

D	b		181	CACACACACACACACACACACACACACACACACACAATAGCGGGCATGTGTTGG	240
O	y		241	GCACTGTAAATCCAGCTACTTGGAGCGCTGAGGCACAGAATGAATTGAACCCAGAGG	300
D	b		241	GCACTGTAAATCCAGCTACTTGGAGCGCTGAGGCACAGAATGAATTGAACCCAGAGG	300
O	y		301	CGGAGTTGCCAGTGAGCTGAGATCATGCCATTGCCACTCCAGCCTGGGTGACGAGTGAAAA	360
D	b		301	CGGAGTTGCCAGTGAGCTGAGATCATGCCATTGCCACTCCAGCCTGGGTGACGAGTGAAAA	360
O	y		361	AAAAATAATGATAATAAAGAGAGCAAGGTGACCACAAAAGAGATAAGCTGGAATAATTT	420
D	b		361	AAAAATAATGATAATAAAGAGAGCAAGGTGACCACAAAAGAGATAAGCTGGAATAATTT	420
O	y		421	GTTAAATGGTGGCCTCTTCTTAAGCTGCATATGTTAAGTTATTTTTCCCTAGT	480
D	b		421	GTTAAATGGTGGCCTCTTCTTAAGCTGCATATGTTAAGTTATTTTTCCCTAGT	480
O	y		481	AGCGAATTCTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAGGTGTAT	540
D	b		481	AGCGAATTCTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAGGTGTAT	540
O	y		541	AACCTATATAGTGAAATAATAGTCCAATTATTCTTTGAAGTAGTAAATATGTAACG	600
D	b		541	AACCTATATAGTGAAATAATAGTCCAATTATTCTTTGAAGTAGTAAATATGTAACG	600
O	y		601	AACTCCTTAAGGCCAGTTGTATACCCAGGGCAAACGCTTCTAACATCTTTATTATCTA	660
D	b		601	AACTCCTTAAGGCCAGTTGTATACCCAGGGCAAACGCTTCTAACATCTTTATTATCTA	660
O	y		661	CGCAGTGGGTAGGAGGTGGGTGAGTGCCCCCTTCCAGCTGATACTGTCAAAACAGGAA	720
D	b		661	CGCAGTGGGTAGGAGGTGGGTGAGTGCCCCCTTCCAGCTGATACTGTGTCAAAACAGGAA	720
O	y		721	GCAAGTTATATCTCTGTTCATAGGAACATGAATAGAGGCCCTTAGTTGTGACTATTAA	780
D	b		721	GCAAGTTATATCTCTGTTCATAGGAACATGAATAGAGGCCCTTAGTTGTGACTATTAA	780
O	y		781	AAAACAAAAAACCTGCCTTAAGAGTTTTCACCTGACTACAAAGTGTAACCTTCTCTGT	840
D	b		781	AAAACAAAAAACCTGCCTTAAGAGTTTTCACCTGACTACAAAGTGTAACCTTCTCTGT	840
O	y		841	GTTTAGAGGAGTGGGTGAGTTTAACTCAGATCCTCTCAATGGGAAAAATAAAAGCCACC	900
D	b		841	GTTTAGAGGAGTGGGTGAGTTTAACTCAGATCCTCTCAATGGGAAAAATAAAAGCCACC	900
O	y		901	AAAAAAAAAAAAAAAAAAAAAACCBAATAACACAGACATCCCAAGTGTGCAGTTGCAAGG	960
D	b		901	AAAAAAAAAAAAAAAAAAAAAACCBAATAACACAGACATCCCAAGTGTGTGCAAGG	960
O	y		961	CTGCTTTGTGTGCCACTTCTCCACATCTTTTCTCATCATCTAAGCAGATGTAGTG	1020
D	b		961	CTGCTTTGTGTGCCACTTCTCCACATCTTTTCTCATCATCTAAGCAGATGTAGTG	1020
O	y		1021	ATGAGCGGCTGGCAGCCACACGTTTCATTGGAAAAAGTCAGATTGGATTGGCCAGG	1080
D	b		1021	ATGAGCGGCTGGCAGCCACACGTTTCATTGGAAAAAGTCAGATTGGATTGGCCAGG	1080
O	y		1081	CATGTAGCTCTCAGGCTTGAAGCGATTACAGGTAAGTTTGTCAACTTGCACGACTCC	1140
D	b		1081	CATGTAGCTCTCAGGCTTGAAGCGATTACAGGTAAGTTTGTCAACTTGCACGACTCC	1140
O	y		1141	CAGCCAGTGAGGTTTCTTAAAGAAACGTATATGAAGACAGGGTCTTTCAATTACAGTT	1197
D	b		1141	CAGCCAGTGAGGTTTCTTAAAGAAACGTATATGAAGACAGGGTCTTTCAATTACAGTT	1197

RESULT 2	.
HS91J24	
LOCUS	HS91J24
DEFINITION	95267 bp DNA linear PRI 18-MAY-2005
	Human DNA sequence from clone RPL-91J24 on chromosome 6q24 Contains
	a pseudogene similar to single-stranded DNA-binding protein, the 5'

ACCESSION	AL024474
VERSION	AL024474.1
KEYWORDS	HTG; CpG island; DMDL; DRP; DRP1; dystrophin; single-stranded DNA-binding protein; UTRN; utrophin.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 95267)
TITLE	Tubby, B.
JOURNAL	Direct Submission Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Aug 5, 1998 this sequence version replaced gi:3355594.
COMMENT	

end of the UTRN gene for utrophin (dystrophin-related protein) (DRP, DRP1, DMDL) and a putative novel gene. Contains a Cpg island, complete sequence.

AL024474.1 GI:3395511
HTG; Cpg island; DMDL; DRP; DRP1; dystrophin; single-stranded DNA-binding protein; UTRN; utrophin.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 95267)
Tubby, B.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Aug 5, 1998 this sequence version replaced gi:3355594.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>
RPL-91J24 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2

RPL-91J24 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality ≥ 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

1.95267

/organism="Homo sapiens"

```
/mol type="genomic DNA"
```

/db xref="RZPD:RPCIP704J2491"

/db_xref="taxon:9606"

```
/chromosome="6"
```

```
/map="q24"
```

```
/clone="RP1-91J24"
```

```
/clone lib="RPCI-1"
```

complement (1577. .2167)

/locus tag="RP1-91J24.1-001"

/pseudo

complement (1577. .2167)

/locus tag="RP1-91J24.1

```
/note="match: proteins:
```

/"acc: procedure
P09381"

10000
/pseudo

```

/codon start=1
/pbcuac

```

join(30244, 30364, 3628

```
join(30244..30303,30260
/gene="UTRN"
```

```

/geom= vrn
/locus tag="RP11-352E13

```

join(30244, 30364, 3628

```
join(50211 : 50202, 50202  
/gene="UTR"
```

```
/locus tag="RP11-352E13
```

```
/product="utrophin (hom
```

```
/procacc - u=oprun: \:con
/note="match: ESTS: BGS
```

join(36372, 36450, AL59

[illegible]

AL590488.13:14057. .14149,AL590488.13:15728. .15805,
AL590488.13:16165. .16257,AL590488.13:18466. .18638,
AL590488.13:20957. .21072,AL590488.13:21760. .21920,
AL590488.13:28114. .28317,AL590488.13:29744. .29925,
AL590488.13:30924. .31074,AL590488.13:32531. .32650,
AL590488.13:36460. .36561,AL590488.13:39390. .39497,
AL590488.13:39781. .39960,AL590488.13:40779. .40948,
AL590488.13:43549. .43672,AL590488.13:45961. .46048,
AL590488.13:50949. .51190,AL590488.13:51353. .51533,
AL590488.13:53376. .53521,AL590488.13:54833. .55045,
AL590488.13:65899. .66012,AL590488.13:66783. .66938,
AL590488.13:71991. .72161,AL590488.13:74388. .74567,
AL590488.13:77564. .77698,AL590488.13:79727. .79876,
AL590488.13:80852. .81013,AL590488.13:82250. .82378,
AL590488.13:83108. .83281,AL590488.13:85480. .85635,
AL590488.13:91436. .91606,AL590488.13:103123. .103302,
AL590488.13:106088. .106216,AL590488.13:106829. .106999,
AL590488.13:108408. .108566,AL590488.13:108990. .109127,
AL590488.13:114159. .114350,AL590488.13:115195. .115367,
AL590488.13:123231. .123381,AL590488.13:125264. .125439,
AL513475.13:2282. .2417,AL513475.13:3994. .4143,
AL513475.13:7495. .7570,AL513475.13:13340. .13554,
AL513475.13:15665. .15782,AL513475.13:19388. .19593,
AL513475.13:21857. .22011,AL513475.13:41799. .41988,
AL513475.13:143106. .143278,AL513475.13:164787. .164943,
AL513184.7:3405. .3534,AL513184.7:21294. .21562,
AL513184.7:24854. .25000,AL513184.7:27768. .27846,
AL513184.7:30977. .31037,AL513184.7:44955. .45016,
AL513184.7:47338. .47412,AL513184.7:54970. .55171,
AL513184.7:62242. .62327,AL513184.7:66882. .67039,
AL513184.7:70872. .71038,AL513184.7:76084. .76195,
AL513184.7:12569. .12705,AL513184.7:19035. .19073,
AL57149.13:19298. .19363,AL57149.13:20477. .20542,
AL57149.13:30860. .30971,AL57149.13:32427. .32519,
AL57149.13:38492. .38514,AL57149.13:42676. .44721)
/gene="UTRN"
/locus_tag="RP11-352E13.1-001"
join(36372. .36450,AL590704.13:54399. .54460,
AL590488.13:14057. .14149,AL590488.13:15728. .15805,
AL590488.13:16165. .16257,AL590488.13:18466. .18638,
AL590488.13:20957. .21072,AL590488.13:21760. .21920,
AL590488.13:28114. .28317,AL590488.13:29744. .29925,
AL590488.13:30924. .31074,AL590488.13:32531. .32650,
AL590488.13:36460. .36561,AL590488.13:39390. .39497,
AL590488.13:39781. .39960,AL590488.13:40779. .40948,
AL590488.13:43549. .43672,AL590488.13:45961. .46048,
AL590488.13:50949. .51190,AL590488.13:51353. .51533,
AL590488.13:53376. .53521,AL590488.13:54833. .55045,
AL590488.13:65899. .66012,AL590488.13:66783. .66938,
AL590488.13:71991. .72161,AL590488.13:74388. .74567,
AL590488.13:77564. .77698,AL590488.13:79727. .79876,
AL590488.13:80852. .81013,AL590488.13:82250. .82378,
AL590488.13:83108. .83281,AL590488.13:85480. .85635,
AL590488.13:91436. .91606,AL590488.13:103123. .103302,
AL590488.13:106088. .106216,AL590488.13:106829. .106999,
AL590488.13:108408. .108566,AL590488.13:108990. .109127,
AL590488.13:114159. .114350,AL590488.13:115195. .115367,
AL590488.13:123231. .123381,AL590488.13:125264. .125439,
AL513475.13:2282. .2417,AL513475.13:3994. .4143,
AL513475.13:7495. .7570,AL513475.13:13340. .13554,
AL513475.13:15665. .15782,AL513475.13:19388. .19593,
AL513475.13:21857. .22011,AL513475.13:41799. .41988,
AL513475.13:143106. .143278,AL513475.13:164787. .164943,
AL513184.7:3405. .3534,AL513184.7:21294. .21562,
AL513184.7:24854. .25000,AL513184.7:27768. .27846,
AL513184.7:30977. .31037,AL513184.7:44955. .45016,
AL513184.7:47338. .47412,AL513184.7:54970. .55171,
AL513184.7:62242. .62327,AL513184.7:66882. .67039,
AL513184.7:70872. .71038,AL513184.7:76084. .76195,
AL513184.7:12569. .12705,AL513184.7:19035. .19073,
AL57149.13:19298. .19363,AL57149.13:20477. .20542,
AL57149.13:30860. .30971,AL57149.13:32427. .32519,
AL57149.13:38492. .38514,AL57149.13:42676. .42684)
/gene="UTRN"
/locus_tag="RP11-352E13.1-001"

CDS

AL57149.13:30860. .30971,AL57149.13:32427. .32519,
AL57149.13:38492. .38514,AL57149.13:42676. .44721)
/gene="UTRN"
/locus_tag="RP11-352E13.1-001"
/product="utrophin (homologous to dystrophin)"
/note="match: ESTs: AA808926 AA984785 AT141950 AL040638
AL046945 AU123569 AW236949 AW244410 AW359744 AW467977
AW837023 AW837028 AW837031 AW937757 BE816112 BE837146
BE837214 BE837283 BE930574 BF076633 BF979631 BG200973
BG212445 BM676900 BM682036 BM970690 BM971046 BQ015824
match: cDNAs: AJ002967 AK020881 X15488 X69086 X83506
Y12229"

join(36372. .36450,AL590704.13:54399. .54460,
AL590488.13:14057. .14149,AL590488.13:15728. .15805,
AL590488.13:16165. .16257,AL590488.13:18466. .18638,
AL590488.13:20957. .21072,AL590488.13:21760. .21920,
AL590488.13:28114. .28317,AL590488.13:29744. .29925,
AL590488.13:30924. .31074,AL590488.13:32531. .32650,
AL590488.13:36460. .36561,AL590488.13:39390. .39497,
AL590488.13:39781. .39960,AL590488.13:40779. .40948,
AL590488.13:43549. .43672,AL590488.13:45961. .46048,
AL590488.13:50949. .51190,AL590488.13:51353. .51533,
AL590488.13:53376. .53521,AL590488.13:54833. .55045,
AL590488.13:65899. .66012,AL590488.13:66783. .66938,
AL590488.13:71991. .72161,AL590488.13:74388. .74567,
AL590488.13:77564. .77698,AL590488.13:79727. .79876,
AL590488.13:80852. .81013,AL590488.13:82250. .82378,
AL590488.13:83108. .83281,AL590488.13:85480. .85635,
AL590488.13:91436. .91606,AL590488.13:103123. .103302,
AL590488.13:106088. .106216,AL590488.13:106829. .106999,
AL590488.13:108408. .108566,AL590488.13:108990. .109127,
AL590488.13:114159. .114350,AL590488.13:115195. .115367,
AL590488.13:123231. .123381,AL590488.13:125264. .125439,
AL513475.13:2282. .2417,AL513475.13:3994. .4143,
AL513475.13:7495. .7570,AL513475.13:13340. .13554,
AL513475.13:15665. .15782,AL513475.13:19388. .19593,
AL513475.13:21857. .22011,AL513475.13:41799. .41988,
AL513475.13:143106. .143278,AL513475.13:164787. .164943,
AL513184.7:3405. .3534,AL513184.7:21294. .21562,
AL513184.7:24854. .25000,AL513184.7:27768. .27846,
AL513184.7:30977. .31037,AL513184.7:44955. .45016,
AL513184.7:47338. .47412,AL513184.7:54970. .55171,
AL513184.7:62242. .62327,AL513184.7:66882. .67039,
AL513184.7:70872. .71038,AL513184.7:76084. .76195,
AL57149.13:12569. .12705,AL57149.13:19035. .19073,
AL57149.13:19298. .19363,AL57149.13:20477. .20542,
AL57149.13:27467. .27625,AL57149.13:27988. .28228,
AL57149.13:30860. .30971,AL57149.13:32427. .32519,
AL57149.13:38492. .38514,AL57149.13:42676. .42684)
/gene="UTRN"
/locus_tag="RP11-352E13.1-001"

Query Match 100.0%; Score 1197; DB 8; Length 95267;
Best Local Similarity 100.0%; Pred. No. 7,le-293;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTATTTCACACAAGCAAGAAAAAGATGAGAGAGACTAGAAAGTATGATGATC 60
DB 87734 TTTCTATTTCACACAAGCAAGAAAAAGATGAGAGAGACTAGAAAGTATGATGATC 87793
QY 61 ATATGAATTAATGATTTTCTGCTTTTTCATGTATGTGTGACATGCAGAGTGAC 120
DB 87794 ATATGAATTAATGATTTTCTGCTTTTTCATGTATGTGTGACATGCAGAGTGAC 87853
QY 121 AGCAGAGTTGAGACCAGCCTGAGCAACATGTTGAATCCCGTCTTAACAACACA 180
DB 87854 AGCAGAGTTGAGACCAGCCTGAGCAACATGTTGAATCCCGTCTTAACAACACA 87913
QY 181 CACACACACACACACACACACACACACACACACACACACACCAATAGCCGGCATGTGTGG 240
DB 87914 CACACACACACACACACACACACACACACACACACACACACCAATAGCCGGCATGTGTGG 87973
QY 241 GCACCTGTAATCCAGCTACTTGGAGGGCTGAGGCACAAGATGACTTGAACCCAGAGG 300

Db 87974 GCACCTGTATCCAGCTACTTGGAGGCTGAGGCACAGAATGACTTGAACCCAGAGG 88033
QY 301 CGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 360
Db 88034 CGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 88093
QY 361 AAAAATAATGATATAAAGAGAGCAAGGTGACCAACAAAAGAGATAGGCTGAAAAATT 420
Db 88094 AAAAATAATGATATAAAGAGAGCAAGGTGACCAACAAAAGAGATAGGCTGAAAAATT 88153
QY 421 GTCTAATGGTGCCCTCTCTCTTATAGCTGCATATGTTAGTTATTTTCCCTAGT 480
Db 88154 GTCTAATGGTGCCCTCTCTCTTATAGCTGCATATGTTAGTTATTTTCCCTAGT 88213
QY 481 AGCGAATCTTAAGGATGAAGAATAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 540
Db 88214 AGCGAATCTTAAGGATGAAGAATAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 88273
QY 541 AACTACTATAGTGAATAATAGTCCATTATTTCTTGAAGTATAGTTAATATGTAACG 600
Db 88274 AACTACTATAGTGAATAATAGTCCATTATTTCTTGAAGTATAGTTAATATGTAACG 88333
QY 601 AACTCCTTAAGGCCAGTTGTATACCCAGGGCAACGCCCTTCTAACATCTTTATTTCTA 660
Db 88334 AACTCCTTAAGGCCAGTTGTATACCCAGGGCAACGCCCTTCTAACATCTTTATTTCTA 88393
QY 661 CGCAGTGGGTAGGAGGTGGGTGAGTGCCTTCCAGCTGATATCTGTCAAAACAGGAA 720
Db 88394 CGCAGTGGGTAGGAGGTGGGTGAGTGCCTTCCAGCTGATATCTGTCAAAACAGGAA 88453
QY 721 GCAAGTTATATCTCTGTATAGGAACATGATATAGAGGCCCTTAGTTGTACTATTTAA 780
Db 88454 GCAAGTTATATCTCTGTATAGGAACATGATATAGAGGCCCTTAGTTGTACTATTTAA 88513
QY 781 AAAACAAAAAACCTGCTTAAGAGTTTTCAGTACTACAAAGTGAATCTTCTCTGTGT 840
Db 88514 AAAACAAAAAACCTGCTTAAGAGTTTTCAGTACTACAAAGTGAATCTTCTCTGTGT 88573
QY 841 GTTTAGAGAGGTGGGTTAGTTAGTCAAGATCCTCTCATGGGAAAAATAAAGCCACC 900
Db 88574 GTTTAGAGAGGTGGGTTAGTTAGTCAAGATCCTCTCATGGGAAAAATAAAGCCACC 88633
QY 901 AAAAAAAAAAAAAAAAAAACCATAATACACAGACATCCCAAGTGTGACGTTGCAAG 960
Db 88634 AAAAAAAAAAAAAAAAAAACCATAATACACAGACATCCCAAGTGTGACGTTGCAAG 88693
QY 961 CTGCTTTTGTGTCCACTTCTCCACATCTTTTCTCATCATCTAAGCAGATGTAGGTG 1020
Db 88694 CTGCTTTTGTGTCCACTTCTCCACATCTTTTCTCATCATCTAAGCAGATGTAGGTG 88753
QY 1021 ATGAGCGGCTGGCAGCCACACGTTTCATTGAAAAAGTGACAGATTGGATTGGCAGGG 1080
Db 88754 ATGAGCGGCTGGCAGCCACACGTTTCATTGAAAAAGTGACAGATTGGATTGGCAGGG 88813
QY 1081 CATGTAGCTCTCCAGGCTTGCAAGCGATTACAGGTAAGTTGTCAACTTGACGACTCC 1140
Db 88814 CATGTAGCTCTCCAGGCTTGCAAGCGATTACAGGTAAGTTGTCAACTTGACGACTCC 88873
QY 1141 CAGCCAGTGAGGTTTCTTAAGAAACGTCTATGAAGACAGGGTCTTTCATTCAGTT 1197
Db 88874 CAGCCAGTGAGGTTTCTTAAGAAACGTCTATGAAGACAGGGTCTTTCATTCAGTT 88930

RESULT 3
AX467737 1465 bp DNA linear PAT 16-JUL-2002
LOCUS AX467737
DEFINITION Sequence 1 from Patent WO0236620.
ACCESSION AX467737
VERSION AX467737.1 GI:21900909
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS 1 Ruegg,M.A. and Brignuet,A.
TITLE Ets-transcription factor related compound specific promoter and transactivators thereof
JOURNAL Patent: WO 0236620-A 1 10-MAY-2002;
FEATURES Myocontract Pharmaceutical Research AG (CH)
source 1. 1465
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 83.9%; Score 1004; DB 6; Length 1465;
Best Local Similarity 100.0%; Pred. No. 8e-244;
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTCTATTTCACACACAGCAAGAAAAAGATGAGAGAGGACTAGAAAGTAGATGTATC 60
Db 462 TTTCTATTTCACACACAGCAAGAAAAAGATGAGAGAGGACTAGAAAGTAGATGTATC 521
QY 61 ATATGAATATGATTTTCTTCTTCTTTTGTGATATGTGTGACACATGACAGATGAC 120
Db 522 ATATGAATATGATTTTCTTCTTCTTTTGTGATATGTGTGACACATGACAGATGAC 581
QY 121 AGCAGAGTTGAGACAGCCTGACCAACATGTTAAATCCCGTCTACTAAACACACA 180
Db 582 AGCAGAGTTGAGACAGCCTGACCAACATGTTAAATCCCGTCTACTAAACACACA 641
QY 181 CACACACACACACACACACACACACACACACACACATAGCCGGGCAATGGTGTGG 240
Db 642 CACACACACACACACACACACACACACACACACACATAGCCGGGCAATGGTGTGG 701
QY 241 GCACCTGTAATCCAGCTACTTGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAG 300
Db 702 GCACCTGTAATCCAGCTACTTGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAG 761
QY 301 CGAGGTTGACGTGAGTCAATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 360
Db 762 CGAGGTTGACGTGAGTCAATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 821
QY 361 AAAAATATGATATATAAGAGAGCAAGGTGACCAAAAAGAGATAGCTGAAAAATT 420
Db 822 AAAAATATGATATATAAGAGAGCAAGGTGACCAAAAAGAGATAGCTGAAAAATT 881
QY 421 GTCTAATAGTGGCTCTTCTCTTATAGCTGCATATGTTAGTTATTTTCCCTAGT 480
Db 882 GTCTAATAGTGGCTCTTCTCTTATAGCTGCATATGTTAGTTATTTTCCCTAGT 941
QY 481 AGCGAATCTTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 540
Db 942 AGCGAATCTTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 1001
QY 541 AACTACTATAGTGAATAATAGTCCCAATTATCTTGAAGTATAGTTAATATGTAACG 600
Db 1002 AACTACTATAGTGAATAATAGTCCCAATTATCTTGAAGTATAGTTAATATGTAACG 1061
QY 601 AAACCTCTAAGGCCAGTTGTATATCCAGGGCAACGCCCTTCTAACATCTTTATCTA 660
Db 1062 AAACCTCTAAGGCCAGTTGTATATCCAGGGCAACGCCCTTCTAACATCTTTATCTA 1121
QY 661 CGCAGTGGGTAGGAGGTGGGTGAGTGCCTTCCAGCTGATATCTGCAAAACAGGAA 720
Db 1122 CGCAGTGGGTAGGAGGTGGGTGAGTGCCTTCCAGCTGATATCTGCAAAACAGGAA 1181
QY 721 GCAAGTTATATCTCTGTATAGGAACATGATAGAGCCCTTAGTTGACTATTAA 780
Db 1182 GCAAGTTATATCTCTGTATAGGAACATGATAGAGCCCTTAGTTGACTATTAA 1241
QY 781 AAAACAAAAAACCTGCTTAAGAGTTTTCAGTACTACAAAGTGAATCTTCTCTGTGT 840

Db 1242 AAAACAAAACCTGCCTAAGAGATTTCAGTACTACAAAGTGTAACTTCTCTGTGT 1301

QY 841 GTTTAGAGAGGTGGGGTTAGTTAGTCAGATCCTCTCATGGGAAAAATAAAGCCACC 900

Db 1302 GTTTAGAGAGGTGGGGTTAGTTAGTCAGATCCTCTCATGGGAAAAATAAAGCCACC 1361

QY 901 AAAAAAAAAAAAAAAAAACCCAAAATAACACAGAGATCCCAAGTGTGCAAGTTCGAAG 960

Db 1362 AAAAAAAAAAAAAAAAAACCCAAAATAACACAGAGATCCCAAGTGTGCAAGTTCGAAG 1421

QY 961 CTGCTTTGTGTGCCACTTCCTCCACATCTTTTCTCATCATC 1004

Db 1422 CTGCTTTGTGTGCCACTTCCTCCACATCTTTTCTCATCATC 1465

RESULT 4
AX541211
LOCUS AX541211 823 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 458 from Patent WO02055700.
ACCESSION AX541211
VERSION AX541211.1 GI:25274637
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Escobedo, J., Garcia, P. D., Kassam, A., Lamson, G., Drmanac, R., Crkvenjakov, R., Dickson, M., Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V., Jones, W. L., Stache-Crain, B. and Scott, E. M.
TITLE Human genes and gene expression products isolated from human prostate
JOURNAL Patent: WO 02055700-A 458 18-JUL-2002;
Chiron Corporation (US) ; HYSEQ, INC. (US)
FEATURES
source 1. 823
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 39.0%; Score 467; DB 6; Length 823;
Best Local Similarity 90.9%; Pred. No. 1e-107;
Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;

QY 1 TTTCTATTTCACAACAAGAAAAAGATGAGAGAGACTAGAAAGTAGATGTGATC 60

Db 145 TTTCTATTTCACAACAAGAAAAAGATGAGAGAGACTAGAAAGTAGATGTGATC 204

QY 61 ATATGAATATGATTTCTTCTTGTGATGATGTGTGACACATGAGAAAGTGAC 120

Db 205 ATATGAATATGATTTCTTCTTGTGATGATGTGTGACACATGAGAAAGTGAC 264

QY 121 AGCAGAGTTGAGACCGCTGACCAACATGATGAAATCCCGTCTTACTAAACACACA 180

Db 265 AGCAGAGTTGAGACCGCTGACCAACATGATGAAATCCCGTCTTACTAAACACACA 324

QY 181 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 240

Db 325 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 384

QY 241 GCACCTGTAATCCAGCTACTTGGAGGCTGAGGCACAGAATGACTTGAACCCAGAGG 300

Db 385 GCACCTGTAATCCAGCTACTTGGAGGCTGAGGCACAGAATGACTTGAACCCAGAGG 444

QY 301 CGAGGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360

Db 445 CGAGGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 504

QY 361 AAAAATATGATAATAAAGAGCAAGGTGACCAAAAAGAGATAGGCTGAAAAATTT 420

Db 505 AAAATAGTATATANAGAGAGCGGGTGACCAAAAAGAGAAATANGNTGAANATTT 564

QY 421 GTCTAAATGGTGG-CCTCTTCTTATAGCTGCATATAGTTAGTTATTTTCCCTAG 479

Db 565 GTCTAANNNGGGGCCCTCTCTCTTATAGCTGCATATAGTGAATTTATTTTCCCTA- 623

QY 480 TAGCGAATCTTAAGGATGAAGAAGAAATCCTTTTACGTTTACTCCCAAGGTGTGA 539

Db 624 NAGCGAATCT-AGGATGAANAAAAAATNCTTTTACGTTTCTCCCAAGGNGTATA 682

QY 540 TAACTACTATAGTGAATAATATAGTCCATTTATCTTTGAA 581

Db 683 CCTCTA---TGGAAATATATAGCCCATTTTCTTTGAGTA 720

RESULT 5
HSA250044
LOCUS HSA250044 458 bp DNA linear PRI 26-NOV-1999
DEFINITION Homo sapiens partial UTRNB gene for utrophin, exon 1 and minimal promoter.
ACCESSION AJ250044
VERSION AJ250044.1 GI:6469369
KEYWORDS UTRNB gene; utrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Burton, E. A., Tinsley, J. M., Holzfeld, P. J., Rodrigues, N. R. and Davies, K. E.
TITLE A second promoter provides an alternative target for therapeutic up-regulation of utrophin in Duchenne muscular dystrophy
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (24), 14025-14030 (1999)
PUBMED 10570192
REFERENCE 2 (bases 1 to 458)
AUTHORS Tinsley, J. M.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-1999) Tinsley J. M., Department of Human Anatomy and Genetics, University of Oxford, South Parks Road, Oxford, OX1 3QX, UNITED KINGDOM

FEATURES
source 1. 458
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

promoter <1..257
/note="minimal"
258..430
/gene="UTRNB"
258..>430
/gene="UTRNB"
258..430
/gene="UTRNB"
/number=1
258..337
/gene="UTRNB"
338..>430
/gene="UTRNB"
/function="unknown"
/codon_start=1
/product="utrophin"
/protein_id="CAB61826.1"
/db_xref="GI:6469370"
/db_xref="UniProt/TREMBL:Q9UJ40"
/translation="MSGIAATTFHWKCKRLDLPGHVALQACKRLP"

ORIGIN

Query Match 38.3%; Score 458; DB 8; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.1e-105;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 684 GAGTGCCCTTCCAGCTGATCTGTCAAAACAGGAAGCAAGTATATCTCTGTGATA 743

Db 1 ||||| GAGTGCCTTCCAGCTGATCTGCAAAAACAGGAAGCAAGTAAATCTCTGTGATA 60

Qy 744 GGAACATGAATAGAGGCCCTTAGTTGCACTATTAAAAAACAACAACTGCCTAAGGA 803

Db 61 GGAACATGAATAGAGGCCCTTAGTTGCACTATTAAAAAACAACAACTGCCTAAGGA 120

Qy 804 GTTTCACCTGACTACAAAGTGTAACTTCTCTCTGTGTTAGAGAGGTGGGTAGGT 863

Db 121 GTTTCACCTGACTACAAAGTGTAACTTCTCTCTGTGTTAGAGAGGTGGGTAGGT 180

Qy 864 TTAGTCAGATCCTCTCATGGGAAAAATAAAGCCACCAAAAAAACC 923

Db 181 TTAGTCAGATCCTCTCATGGGAAAAATAAAGCCACCAAAAAAACC 240

Qy 924 AAAATAACACAGACATCCAGTGTGAGTTGCAAGGCTGTTTGTCTCCACTTCTC 983

Db 241 AAAATAACACAGACATCCAGTGTGAGTTGCAAGGCTGTTTGTCTCCACTTCTC 300

Qy 984 CACATCTTTTCTCTCATCTAAGCAGATGTAGTGAGGCGGCTGCGAGCCACCAC 1043

Db 301 CACATCTTTTCTCTCATCTAAGCAGATGTAGTGAGGCGGCTGCGAGCCACCAC 360

Qy 1044 GTTTCATGGAAGGATGAGTTGCACTTGCCAGGCGATGCTCTCCAGGCTTGCAA 1103

Db 361 GTTTCATGGAAGGATGAGTTGCACTTGCCAGGCGATGCTCTCCAGGCTTGCAA 420

Qy 1104 GCGATTACAGTAAGTTGTCACTTGACAGACTCCC 1141

Db 421 GCGATTACAGTAAGTTGTCACTTGACAGACTCCC 458

RESULT 6
AC156852/c 224348 bp DNA linear HTG 01-JUL-2005
LOCUS Bos taurus clone CH240-59D8, *** SEQUENCING IN PROGRESS ***, 34
DEFINITION unordered pieces.

ACCESSION AC156852
VERSION AC156852.2 GI:68303076
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Muzny,D.Marie, Metzker,M.Lee, Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalio,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Liu,J., Liu,J., Liu,W., Liu,Y., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Lorensuhewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackeleme,O., Okwuon,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 29, 2005 this sequence version replaced gi:58652209. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: FDSO
Center clone name: CH240-59D8

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 210704 bases at least Q40
Consensus quality: 213145 bases at least Q30
Consensus quality: 215438 bases at least Q20
Estimated insert size: 211859; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 1621: contig of 1621 bp in length
* 1622 2987: gap of 1366 bp
* 2988 4420: contig of 1433 bp in length
* 4421 4520: gap of unknown length
* 4521 10376: contig of 5856 bp in length
* 10377 11040: gap of 664 bp
* 11041 13556: contig of 2516 bp in length
* 13557 13606: gap of 50 bp
* 13607 15408: contig of 1802 bp in length
* 15409 15887: gap of 479 bp
* 15888 17253: contig of 1366 bp in length
* 17254 17864: gap of 611 bp
* 17865 25099: contig of 7235 bp in length
* 25100 25149: gap of 50 bp
* 25150 31522: contig of 6373 bp in length
* 31523 31918: gap of 396 bp
* 31919 35351: contig of 3433 bp in length
* 35352 35401: gap of 50 bp
* 35402 40574: contig of 5173 bp in length
* 40575 40668: gap of 94 bp
* 40669 48691: contig of 8023 bp in length
* 48692 48741: gap of 50 bp
* 48742 53109: contig of 4368 bp in length
* 53110 53159: gap of 50 bp
* 53160 61668: contig of 8509 bp in length
* 61669 61718: gap of 50 bp
* 61719 75464: contig of 13746 bp in length
* 75465 75514: gap of 50 bp
* 75515 77844: contig of 2330 bp in length
* 77845 77944: gap of unknown length
* 77945 85565: contig of 7621 bp in length
* 85566 85615: gap of 50 bp
* 85616 107375: contig of 21760 bp in length
* 107376 107425: gap of 50 bp
* 107426 111000: contig of 3575 bp in length
* 111001 111050: gap of 50 bp
* 111051 117345: contig of 6295 bp in length
* 117346 117395: gap of 50 bp
* 117396 125897: contig of 8502 bp in length
* 125898 125947: gap of 50 bp
* 125948 127099: contig of 1152 bp in length
* 127100 127449: gap of 350 bp
* 127450 153169: contig of 25720 bp in length
* 153170 153219: gap of 50 bp
* 153220 185179: contig of 31960 bp in length
* 185180 185229: gap of 50 bp
* 185230 196889: contig of 11660 bp in length
* 196890 196939: gap of 50 bp
* 196940 198078: contig of 1139 bp in length
* 198079 198128: gap of 50 bp
* 198129 209346: contig of 11218 bp in length
* 209347 209396: gap of 50 bp
* 209397 211978: contig of 2582 bp in length
* 211979 212078: gap of unknown length
* 212079 213092: contig of 1014 bp in length
* 213093 213192: gap of unknown length
* 213193 214200: contig of 1008 bp in length
* 214201 214300: gap of unknown length
* 214301 215319: contig of 1019 bp in length
* 215320 215419: gap of unknown length
* 215420 216456: contig of 1037 bp in length
* 216457 216556: gap of unknown length
* 216557 218320: contig of 1764 bp in length
* 218321 218420: gap of unknown length
* 218421 220493: contig of 2073 bp in length
* 220494 220593: gap of unknown length
* 220594 224348: contig of 3755 bp in length.

FEATURES

source
1. .224348
/organism="Bos taurus"
/mol_type="genomic DNA"

/db_xref="taxon:9913"
/clone="CH240-59D8"
1622. .2987
/estimated_length=1366
4421. .4520
/estimated_length=unknown
10377. .11040
/estimated_length=664
13557. .13606
/estimated_length=50
15409. .15887
/estimated_length=479
17254. .17864
/estimated_length=611
25100. .25149
gap

Query Match 29.0%; Score 347.6; DB 14; Length 224348;
Best Local Similarity 74.7%; Pred. No. 1.2e-77;
Matches 589; Conservative 0; Mismatches 144; Indels 55; Gaps 10;

QY 376 AAAGAGCAAGGTGACCAAAAAGATAGCTGAAAA-ATTGCTAAATGTGCG 434
DB 116376 AAAGAGCGAGCAAAATACACAAGATAGAACAGCTGAAAAATATGATCAAACTGTAC 116317
QY 435 CTCTCTCTATAGTCATATGTTAAGTTATTTTCCC-TAGTACGAATTCTAAG 493
DB 116316 CTCTGATCTTATGATCTGTGTGTTGAATGATTTTTCCTAATAAGCATCTCTAAG 116257
QY 494 GGATGAAGAGAAATCTTTTCAGTTTAA-CTTCCCAAGGTGTATACTACTATAGT 552
DB 116256 GAATGAAGAGAAATCTTTTGAAGTTTTCCTCAAGGTGTATATGATATTTGAAGT 116197
QY 553 GAAATATAAGTCCAAATTATTTTGAAGTATATGTAACGAAACTCTAAGG 612
DB 116196 GAAATATAAGTCAATTATTTTGAAGTATATGTAACGAAACTCTCCGT 116137
QY 613 CCAGTTGTATACCCAGGCGCAACGCTTCTAATCACTTTATTTATCTACGAGTGGTAG 672
DB 116136 CCAGTTGTATACCCAGGCGCAACGCTTCTAATCACTTTATTTATGCAAT- -GTGGGTAA 116088
QY 673 GGAGTGGGTGAGTCCCTTCCAGCTGATACGTGC-AAAAAGAGCAAGGTATTA 731
DB 116087 GGAGTGGGTGAGTGTCTTATCACTGACGTGTCAAAAAAGAGTATTA 116028
QY 732 ATCTCTGTCATAGAAACATGAA- -TAGAGCCCTTAGTTGTGACTATTAATAAACAA 786
DB 116027 ATCTCTTGTAGTGTGATCAGAACATTCGAAGGCCCTCTTGTGACAATTAAGAACA 115968
QY 787 -AAAACTGCTTAAGAGATTTTCACTGACTACAAAGTAACTCTCTGTG 841
DB 115967 GTGACAGAAATGACCTTAAGGAATTTTCACTAATCTGGAAGTAACTCTGTG 115908
QY 842 TTTAGAGGAGTGGGTTAGG-TTTAGTCAAGATCTCTCATGGAAAAATTAAGCCACC 900
DB 115907 TTTAGAGGAGATTCAGTTAGGTTTAGTCAAGATCTCTCATGGAAAAATTAAGCCACG 115848
QY 901 AAAAAAAAAAAAAAAAAAACCATAACAGAGACATCCAGTGTGAGTGAAGG 960
DB 115847 GCGAA- -ACAGAGATCCCAAGTGTGAGTGTGAGG 115814
QY 961 CTGCTTTTGT- -GTCCACTTCTCCACATCTTTTCTCATCATTAAGAGATGTAG 1017
DB 115813 CTGCTTCTGTTAAATCCACTTCTGTGATCTTTTCTCTGCTCACTTAAGCAATGTAG 115754
QY 1018 GTGATGAGCGGCTGCGACCAACAGCTTTCAATTGAAAAAGTGAGATTGGATTTGCCA 1077
DB 115753 GTGATGAGCGGCGCGCGCCACCGCGTTTCGTTGAAAAAGTACAGATTGATGTGCCA 115694
QY 1078 GGGCATGTAGCTCTCAGGCTTGCAAGCAATTAACAGTAAGTTGTCACTTGCAAGCAC 1137
DB 115693 GGACACGCAAGCTCGCCAGTGTGCAAGCAATTAACAGTAAGTTGTGCACTTGCCCTGAC 115634
QY 1138 TCCAGGCC 1145

Db 115633 TCCCTGCC 115626

|||||

RESULT 7
AC126325/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RPl1-378H8 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC126325
AC126325 2 GI:22004321
VERSION
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
TITLE Birren,B., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 11, clone RPl1-378H8
REFERENCE
AUTHORS Unpublished
2 (bases 1 to 88251)

TITLE Direct Submission

JOURNAL Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome

REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA

AUTHORS 3 (bases 1 to 88251)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kelis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tessfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome

COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 30, 2002 this sequence version replaced gi:21699256.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: 127750
Center clone name: 378_H_8

* NOTE: This record contains 80 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
1008 1007: contig of 1007 bp in length
1108 1107: gap of 100 bp
2093 2092: contig of 985 bp in length
2193 2192: gap of 100 bp
3207 3206: contig of 1014 bp in length
3307 3306: gap of 100 bp
4320 4319: contig of 1013 bp in length
4420 4419: gap of 100 bp
5427 5426: contig of 1007 bp in length
5527 5526: gap of 100 bp
6529 6528: contig of 1002 bp in length
6629 6628: gap of 100 bp
7629 7628: contig of 1000 bp in length
7729 7728: gap of 100 bp
8682 8681: contig of 953 bp in length
8782 8781: gap of 100 bp
9785 9784: contig of 1003 bp in length
9885 9884: gap of 100 bp
10910 10909: contig of 1025 bp in length
11010 11009: gap of 100 bp
12048 12047: contig of 1038 bp in length
12148 12147: gap of 100 bp
13121 13120: contig of 973 bp in length
13221 13220: gap of 100 bp
14253 14252: contig of 1032 bp in length
14353 14352: gap of 100 bp
15376 15375: contig of 1023 bp in length
15476 15475: gap of 100 bp
16500 16499: contig of 1024 bp in length
17610 17609: gap of 100 bp
17710 17709: contig of 1010 bp in length
18735 18734: contig of 925 bp in length
18735 18734: gap of 100 bp
19730 19729: contig of 995 bp in length
19830 19829: gap of 100 bp
20849 20848: contig of 1019 bp in length
20949 20948: gap of 100 bp
21974 21973: contig of 1025 bp in length
22074 22073: gap of 100 bp
23072 23071: contig of 998 bp in length
23172 23171: gap of 100 bp
24208 24207: contig of 1036 bp in length
24308 24307: gap of 100 bp
25349 25348: contig of 1041 bp in length
25449 25448: gap of 100 bp
26484 26483: contig of 1035 bp in length
26584 26583: gap of 100 bp
27619 27618: contig of 1035 bp in length
27719 27718: gap of 100 bp
28706 28705: contig of 987 bp in length
28806 28805: gap of 100 bp
29806 29805: contig of 1000 bp in length
30921 30920: gap of 100 bp
31021 31020: contig of 1015 bp in length
32029 32028: contig of 1009 bp in length

32030	32129: gap of 100 bp
32130	33108: contig of 979 bp in length
33109	33208: gap of 100 bp
33209	34220: contig of 1012 bp in length
34221	34320: gap of 100 bp
34321	35318: contig of 998 bp in length
35319	35418: gap of 100 bp
35419	36453: contig of 1035 bp in length
36454	36553: gap of 100 bp
36554	37589: contig of 1036 bp in length
37590	37689: gap of 100 bp
37690	38667: contig of 978 bp in length
38668	38767: gap of 100 bp
38768	39782: contig of 1015 bp in length
39783	39882: gap of 100 bp
39883	40879: contig of 997 bp in length
40880	40979: gap of 100 bp
40980	41986: contig of 1007 bp in length
41987	42086: gap of 100 bp
42087	43070: contig of 984 bp in length
43071	43170: gap of 100 bp
43171	44172: contig of 1002 bp in length
44173	44272: gap of 100 bp
44273	45257: contig of 985 bp in length
45258	45357: gap of 100 bp
45358	46358: contig of 1001 bp in length
46359	46458: gap of 100 bp
46459	47461: contig of 1003 bp in length
47462	47561: gap of 100 bp
47562	48600: contig of 1039 bp in length
48601	48700: gap of 100 bp
48701	49738: contig of 1038 bp in length
49739	49838: gap of 100 bp
49839	50838: contig of 1000 bp in length
50839	50938: gap of 100 bp
50939	51968: contig of 1030 bp in length
51969	52068: gap of 100 bp
52069	53071: contig of 1003 bp in length
53072	53171: gap of 100 bp
53172	54141: contig of 970 bp in length
54142	54241: gap of 100 bp
54242	55242: contig of 1001 bp in length
55243	55342: gap of 100 bp
55343	56325: contig of 983 bp in length
56326	56425: gap of 100 bp
56426	57352: contig of 927 bp in length
57353	57452: gap of 100 bp
57453	58477: contig of 1025 bp in length
58478	58577: gap of 100 bp
58578	59577: contig of 1000 bp in length
59578	59677: gap of 100 bp
59678	60706: contig of 1029 bp in length
60707	60806: gap of 100 bp
60807	61844: contig of 1038 bp in length
61845	61944: gap of 100 bp
61945	62948: contig of 1004 bp in length
62949	63048: gap of 100 bp
63049	64072: contig of 1024 bp in length
64073	64172: gap of 100 bp
64173	65155: contig of 983 bp in length

Query Match 22.1%; Score 264.6; DB 14; Length 88251;
Best Local Similarity 91.8%; Pred. No. 1.4e-56;
Matches 367; Conservative 0; Mismatches 22; Indels 11; Gaps 8;

Qy	807	TTCACTGACACAAAGTAACTTCTCTCTGTGTTTAAAGAGAGGTGGGTTAGGTTTA	866
Db	62948	TTCCCTGACTACAAAGGTAAC-TCCCTCTCTGTGTTTAAAGAGAGGTGGGTTTAAAGTTTA	62890
Qy	867	GTGAGATCCTCTCATGCGAAAAATAAAGCCCAAAAAA-----CC	922
Db	62889	GTGAGATCCTCTCTCAAGGAAAAATAAAGCCCAAAAAA-----CC	62830

Qy	923	CAAAATACACAGACATCCAGTGTGAGTTGAAAGCTGTTTGTGTCACCTTCT	982
Db	62829	AAAATNACCACAGACATCCAGTGTGAG-TCGAAGCTGCTTGTGTTCACTTCT	62771
Qy	983	CCACAT-CTTTTCTCATCATCTAAGCAGATGTAGGTATGAGCGGCTGGCAGCCACC	1041
Db	62770	CCACATNCTTTTNTCATCATCTAAGCAGATGTAGGTATGAGCGGCTGGCAGCCACC	62711
Qy	1042	ACGTTTCATT-GGAAAAAGTGACAGATTGGATTTGCCAGGGCATGTAGCTTCCAGGCTTG	1100
Db	62710	ACGTTTCATTGGGAAAAAGTGACAGATTGGATTTGCCAGGGCATGTAGCTTCCAGGCTTG	62651
Qy	1101	CAAGCATTAACAGGTAAGTTGTCAACTTGACGAC-TCCCAGCCAGT-GAGGTTTCT	1158
Db	62650	CAAGCATTAACAGGTAAGTTGTCAACTTGACGACTTCCCAGCCAGTAAAGGTTTCT	62591
Qy	1159	TAAGAAAGCTCTATGAG-ACAGGTTCTTTCATTCACTT	1197
Db	62590	TAAGAAAGCTCTATGAGAACAGAGGTTCTTTCATTCACTT	62551

RESULT 8
AC098332_4/c
WPCOMMENT

Sequence split into 5 fragments LOCUS AC098332 Accession AC098332

Fragment Name	Begin	End
AC098332_0	1	110000
AC098332_1	100001	210000
AC098332_2	200001	310000
AC098332_3	300001	410000
AC098332_4	400001	450102

Continuation (5 of 5) of AC098332 from base 400001 (AC098332 Rattus norvegicus clone C

Query Match 16.3%; Score 195.2; DB 14; Length 50102;
Best Local Similarity 78.1%; Pred. No. 5.7e-39;
Matches 275; Conservative 0; Mismatches 68; Indels 9; Gaps 3;

Qy	783	AACAAAAACCTGCTAAGAGTTTCACTGACTACAAAGTAACTTCTCTGTGTT	842
Db	46093	AAGACCAAAAGTACCTCAAGAGTTTGTACTGACTGACAGATTAACTTCTGTGTTAA	46034
Qy	843	TTAGAGAGGTGGGTTAGGTTTGTACAGATCTCTCATGCGAAAAATAAAGCCAC---	899
Db	46033	TGGGGGAGGTGAGTTAGATTTAATGACAGATCTCTGTGCGAAAAATAAAGGACTCN	45974
Qy	900	-CAAAAAAAAAAAAAAAAACCCAAAATTAACAGACATCCAGTGTGAGTTGAA	958
Db	45973	AAAAAAAAAAAAAAAAAACCAACCACTTAACAGATCCAGTGTGAGTTGCA	45914
Qy	959	GGCTGCTTTTGTGTCACCTTCTCCACATCT-----TTTCTCATCATCTAAGCAGATG	1014
Db	45913	GGCAGCTTTTG-TGTTGACTTCTTCAAGTTACAGTTTCCCTCAACTCAGCCGCCAACG	45855
Qy	1015	TAGGTGATGAGCGGCTGGCAGCCACCAAGTTTCAATTGAAAAAGTGACAGATTGATTTG	1074
Db	45854	TAGGTGATGAGCAGACTGGCAGCCACCAATTTGGTGAAAAAGTGAGGCTGGAATTTG	45795
Qy	1075	CCAGGCGATGATGCTTCCAGGCTTGAAGCGATTACAGGTAAGTTGTCA	1126
Db	45794	CCTGGGCGAGTGCTCTCCAGGCTTGCAGAGATCCCAAGTAAAGTTGTCA	45743

RESULT 9
AC099231 236094 bp DNA linear HTG 10-MAY-2003
LOCUS AC099231
DEFINITION Rattus norvegicus clone CH230-30P24, *** SEQUENCING IN PROGRESS
ACCESSION AC099231
VERSION AC099231.6 GI:30521887
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus.

1 (bases 1 to 236094)

Muzny,D,Marie, Metzker,M,lee, Abramzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Derramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorensuhewa,L, Louiseged,H, Lozado,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwaokelemeh,O, Okwuonu,G, Olarnpunsagoon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfankoch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L,-L, Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savary,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D, Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valase,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wiczysk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R,A.

TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

Direct Submission
2 (bases 1 to 236094)

Worley,K,C.

Direct Submission

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236094)

Rat Genome Sequencing Consortium.

REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:23664925. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHEB
Center clone name: CH230-30P24

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 190957 bases at least Q40
Consensus quality: 195558 bases at least Q30
Consensus quality: 198660 bases at least Q20
Estimated insert size: 199081; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 186058: contig of 186058 bp in length
* 186158: gap of unknown length
* 186159 232369: contig of 46211 bp in length
* 232370 232469: gap of unknown length
* 232470 233616: contig of 1147 bp in length
* 233617 233716: gap of unknown length
* 233717 234819: contig of 1103 bp in length
* 234820 234919: gap of unknown length
* 234920 236094: contig of 1175 bp in length.

FEATURES

source

1. .236094
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-30P24"
4091. .4904
/note="clone boundary
clone end:5p6
site:EcoRI

misc_feature

end sequence: BH294551"
65608. .67141
/note="wgs contig"
184734. .186058
/note="wgs contig"
186059. .186158

misc_feature

/estimated_length=unknown
186159. .187948
/note="wgs contig"
229404. .230840
/note="wgs contig"
232370. .232469

misc_feature

/estimated_length=unknown
233617. .233716
/estimated_length=unknown
234820. .234919
/estimated_length=unknown

misc_feature

gap
gap
gap
gap
ORIGIN

Query Match 16.3%; Score 195.2; DB 14; Length 236094;
Best Local Similarity 78.1%; Pred. No. 5,1e-39;
Matches 275; Conservative 0; Mismatches 68; Indels 9; Gaps 3;

QY 783 AACAAAAACCTGCCTAGAGATTTCACCTGACTACAAAGTGAATCTCTCTGTGT 842
Db 80794 AAGACCAAAAGTACCTCAAGAGTTTGTACTGACTGACAGATTAACTTCTGTGTGTA 80853
QY 843 TTAGAGAGAGTGGGGTTAGTTTGTAGTCAAGTCTCTCATGGGAAAATAAAG----CCA 898
Db 80854 TGGGGGGAGGTGAGTTAGATTAGTCAAGTCTCTGTGGGAAAATAAAGGAGATCA 80913
QY 899 CCAAAAAAAAAAAAAAAAAAAAAAACCCTAATACACAGACATCCAGTGTGAGTTGCA 958
Db 80914 AAAAAAAAAAAAAAAAAAAAAAACAACCACTAACAGACATCCAGTGTGAGTTGCA 80973
QY 959 GGCTGCTTTTGTGTCTCACTTCTCCACATCT---TTTCTCATCATCTAAGCAGATG 1014
Db 80974 GGCAAGCTTTTG-TGTGACTTCTTCAAGTTCAGTTTCTCTCAACTCAGCCGCCAAG 81032
QY 1015 TAGGTGATGAGCGCGCTGGCAGCCACCAAGTTCATTGGAATAAAGTGACAGATTGTTG 1074
Db 81033 TAGGTGATGAGCAGACTGGCAGCCACCAAGTTCATTGGAATAAAGTGAGGCTGATTTG 81092
QY 1075 CCAGGCGATGATGCTTCTCCAGGCTTGCAAGCGATTACCAAGTGAATTGTCA 1126
Db 81093 CCTGGGCGAGGTGTCTTCTCCAGGCTTGCAAGGATCCCAAGTGAATTGTCA 81144

RESULT 10
AC137287/c 247698 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
DEFINITION *** 2 unordered pieces.
AC137287
AC137287.1 GI:25138355
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguitano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemeleh,O., Okwouu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 247698)
Rat Genome Sequencing Consortium.
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KZYL
Center clone name: CH230-unknown
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 214777 bases at least Q40
Consensus quality: 218817 bases at least Q30
Consensus quality: 221608 bases at least Q20
Estimated insert size: 227456; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 238424: contig of 238424 bp in length
* 238425 238524: gap of unknown length
* 238525 247698: contig of 9174 bp in length.
----- Location/Qualifiers
1. 247698
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-unknown"
5857. .8650
/note="wgs_contig"

FEATURES

source

misc_feature

gap	20788. .20887	/estimated_length=unknown
gap	25606. .25705	/estimated_length=unknown
gap	31300. .31399	/estimated_length=unknown
gap	38280. .38379	/estimated_length=unknown
gap	43286. .43385	/estimated_length=unknown
gap	48445. .48544	/estimated_length=unknown
gap	55440. .55539	/estimated_length=unknown
gap	62535. .62634	/estimated_length=unknown
gap	68290. .68389	/estimated_length=unknown
gap	77575. .77674	/estimated_length=unknown
gap	86200. .86299	/estimated_length=unknown
gap	97896. .97995	/estimated_length=unknown
gap	108547. .108646	/estimated_length=unknown
gap	120856. .120955	/estimated_length=unknown
gap	137386. .137485	/estimated_length=unknown
ORIGIN		

Query Match	Best Local Similarity	Score	DB	Length
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;	83.6%;	188.2;	14;	150829;
QY 117	TGACAGCAGAGAGTTGAGAGCCAGCCTGACCAACATGTTGAAATCCCGTCTACTAAACA	176		
Db 140621	TGAGGTCAAGAGTTCAAGAGCCAGCCTGGCCAAATGTTGAAACCCCATCTCTACTAAAAA	1405622		
QY 177	CACACACACACACACACACACACACACACACACACACACACACACATAGCCGCGCATGTG	236		
Db 140561	TACACACACACACACACACACACACACACACACACACACACACAAATTAGCTGGGCGTGTG	140502		
QY 237	GTGGGCACCTGTAATCCAGCTACTTTGGGAGGCTGAGGCACAAGAATGACTTGAACCCAG	296		
Db 140501	GCGAGCACTTGCAATCCCACTACTT-GGAGGCTGGGGCAGAAGATCACTTGAACCCGG	1404433		
QY 297	GAGCGGAGGTTGCAGTGAAGTGAATCATGCCATTGCCACTCCAGCCTGGGTGACGAGTG	356		
Db 140442	GAGGCAAGGTTGCAGTGAAGCGCGATCGTGCCTATTCACCTCCAGCCTGGGCAACAAGAG	1403833		
QY 357	AAAAAAAATATGATTAATTAAGAGAGCA	385		
Db 140382	CAAAACTCTGTCTCAAAAAAAAATTCA	140354		

RESULT 12
 AC016927/c
 LOCUS
 DEFINITION
 AC016927 166206 bp DNA linear HTG 06-DEC-2001
 Homo sapiens chromosome 16 clone RP11-71L14, WORKING DRAFT
 SEQUENCE, 9 unordered pieces.
 AC016927
 AC016927.9 GI:15963639
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 166206)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Albrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL
COMMENT

COMMENT

Benton, J., Bimarge, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Louiseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Rolfe, M.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N.,
Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 166206)

Worley, K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 5, 2001 this sequence version replaced gi:9719614.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----
Center project name: HMSU
Center clone name: RPL1-71L14

----- Summary Statistics -----
Sequencing vector: M13; L08821
Assembly program: Phrap; version 0.990329
Consensus quality: 142794 bases at least Q40
Consensus quality: 155920 bases at least Q30
Consensus quality: 160533 bases at least Q20
Estimated insert size: 165004; sum-of-contigs estimation
Estimated insert size: 176667; agarose-fp estimation
Quality coverage: 3.7x in Q20 bases; agarose-fp estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.
1 40609: contig of 40609 bp in length
* 40610 40709: gap of unknown length
* 40710 76862: contig of 36153 bp in length
* 76863 76962: gap of unknown length
* 76963 105735: contig of 28773 bp in length
* 105736 105835: gap of unknown length
* 105836 130364: contig of 24529 bp in length
* 130365 130464: gap of unknown length
* 130465 144385: contig of 13921 bp in length
* 144386 144485: gap of unknown length
* 144486 152979: contig of 8494 bp in length
* 152980 153079: gap of unknown length
* 153080 159181: contig of 6102 bp in length
* 159182 159281: gap of unknown length
* 159282 162040: contig of 2759 bp in length
* 162041 162140: gap of unknown length
* 162141 166206: contig of 4066 bp in length.

FEATURES
source
1. 166206
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-71L14"
40610. 40709
/estimated_length=unknown
76863. 76962
/estimated_length=unknown
105736. 105835
/estimated_length=unknown
130365. 130464
/estimated_length=unknown
144386. 144485
/estimated_length=unknown
152980. 153079
/estimated_length=unknown
159182. 159281
/estimated_length=unknown
162041. 162140
/estimated_length=unknown

gap
gap
gap
gap
gap
gap
gap
gap
ORIGIN
Query Match 15.7%; Score:188.2; DB 14; Length 166206;
Best Local Similarity 83.6%; Pred. No.3.1e-37;
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 117 TGACAGCAGGAGTTGAGACCAAGCTGACCAATGTTGAATCCCGTCTCTACTAACA 176
Db 129741 TGAGGTCAAGAGTTCAAGACCAGCTGGCCAAACATGTTGAACCCCATCTCTACTAAAA 129682
QY 177 CACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTG 236
Db 129681 TACACACACACACACACACACACACACACACACACAAATTAGCTGGCGGTGTG 129622
QY 237 GTGGGACCTGTAATCCAGCTTGGGAGGCTGAGGCACAGAATGACTTGAACCCAG 296
Db 129621 GCGGACACTGCAATCCAGCTACTT-GGAGGCTGGGGCAGAGAATCACTTGAACCCGG 129563
QY 297 GAGGCGGAGGTTGACGTGAGCTGACATGATCCATTGCACTTCAGCCTGGGTGACGAGTG 356
Db 129562 GAGGCGAGGTTGACGTGAGCCGCGATCGTGCATTGCACTTCAGCCTGGGCAACAAGAG 129503
QY 357 AAAAAAATATGATATAAAGAGCA 385
Db 129502 CAAAACTCTGTCTCAAAAAAATTCA 129474

RESULT 13
AC024903/c 167238 bp DNA linear PRI 06-DEC-2001
LOCUS AC024903 Homo sapiens 16 BAC RP11-71L14 (Roswell Park Cancer Institute Human
DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

BAC Library) complete sequence.
AC024903
AC024903.17 GI:16506331
HTG.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.

1 (bases 1 to 167238)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M.,
Holloway,C., Hollins,B., Homsif., Howard,S., Huber,J., Hulik,S.,
Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojaas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,B.,
Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vaequez,L.,
Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 167238)
Worley,K.C.
Direct Submission
Submitted (02-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (27-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (06-DEC-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 27, 2001 this sequence version replaced gi:15628134.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

QY 357 AAAAAAAAAATGATATTAAGAGCA 385
Db 166375 CAAACTCTGTCTCAAAAAAAAAAATTC 166347

RESULT 14
AC099521/c 167245 bp DNA linear PRI 06-DEC-2001
LOCUS Homo sapiens chromosome 16 clone RP11-71L14, complete sequence.
DEFINITION AC099521
AC099521
VERSION AC099521.2 GI:17386261
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 167245)
TITL DOE Joint Genome Institute.
AUTHORS Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167245)
TITL DOE Joint Genome Institute.
AUTHORS Direct Submission
JOURNAL Submitted (15-NOV-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 167245)
REFERENCE DOE Joint Genome Institute.
AUTHORS Direct Submission
JOURNAL Submitted (06-DEC-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Dec 6, 2001 this sequence version replaced gi:16930937.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source
1. .167245
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-71L14"

ORIGIN

Query Match 15.7%; Score 188.2; DB 8; Length 167245;
Best Local Similarity 83.6%; Pred. No. 3.1e-37;
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
QY 117 TGACAGCAGAGTTCGAGACCAAGCTGACCAATGCCGTCTCTACTAACA 176
Db 166620 TGAGGTCAAGAGTCAAGACCAAGCTGACCAATGCCGTCTCTACTAACA 166561
QY 177 CACACACACACACACACACACACACACACACACATAGCCGGGCATGTTG 236
Db 166560 TACACACACACACACACACACACACACACACACAAATTAGCTGGGCGTGTG 166501

QY 237 GTGGCAGCTGTATATCCCAAGTCTGGAGGCTGAGGCAAGAATGACTTGAACCCAG 296
Db 166500 GCAGACCTGCAATCCCAAGTACTT-GGAGGCTGGGGCAGAAATCACTTGAACCCG 166442
QY 297 GAGCGAGGTTGACAGTGAATCATGCCATTGCACTCCAGCTGGTGACGAGTG 356
Db 166441 GAGCAGAGGTTGACAGTGAAGCGGATCGTGCCATTGCACTCCAGCTGGCAACAAGAG 166382
QY 357 AAAAAAAAAATGATATTAAGAGCA 385
Db 166381 CAAACTCTGTCTCAAAAAAAAAAATTC 166353

RESULT 15
AC009162 181650 bp DNA linear PRI 29-MAR-2003
LOCUS Homo sapiens chromosome 16 clone RP11-76H6, complete sequence.
DEFINITION AC009162
AC009162
VERSION AC009162.11 GI:29366937
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 181650)
TITL DOE Joint Genome Institute, Stanford Human Genome Center and Los
AUTHORS Alamos National Laboratory.
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 181650)
TITL DOE Joint Genome Institute.
AUTHORS Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 181650)
REFERENCE DOE Joint Genome Institute.
AUTHORS Direct Submission
JOURNAL Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 181650)
REFERENCE DOE Joint Genome Institute, Stanford Human Genome Center and Los
AUTHORS Alamos National Laboratory.
TITL Direct Submission
JOURNAL Submitted (29-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Mar 29, 2003 this sequence version replaced gi:15193329.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.shgc.stanford.edu
Quality: Phrap.Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 1.2.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source
1. .181650
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-76H6"

ORIGIN

Query Match 15.7%; Score 188.2; DB 8; Length 181650;
Best Local Similarity 83.6%; Pred. No. 3.1e-37;
Matches 225; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
QY 117 TGACAGCAGAGTTCGAGACCAAGCTGACCAATGCCGTCTCTACTAACA 176
Db 56483 TGAGGTCAAGAGTCAAGACCAAGCTGACCAATGCCGTCTCTACTAACA 56542
QY 177 CACACACACACACACACACACACACACACACACATAGCCGGGCATGTTG 236
Db 166560 TACACACACACACACACACACACACACACACACAAATTAGCTGGGCGTGTG 166501

Db 56543 TACACACACACACACACACACACACACACACACAAATTAGCTGGCGGTGTG 56602
QY 237 GTGGCACCCTGTATATCCCACTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAG 296
Db 56603 GCGAGCACTGCAATCCAGCTACTT-GGAGGCTGGGCAAGAATCACTTGAACCCGG 56661
QY 297 GAGCGGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTG 356
Db 56662 GAGGCAGAGGTTGACGTGAGCCGCGATCGTGCCATTGCACTCCAGCCTGGGCAACAGAG 56721
QY 357 AAAAAAAAAATATGATATAAAGAGAGCA 385
Db 56722 CAAAACTGTGTCTCAAAAAAAATTCA 56750

Search completed: April 18, 2006, 22:13:15
Job time : 6073 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 20:10:38 ; Search time 809 Seconds
(without alignments)
9861.109 Million cell updates/sec

Title: US-10-089-928-1
Perfect score: 1197
Sequence: 1 ttctctattcacaacaagca.....cagggtcttcttcattcagtt 1197

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_21:*
- 1: geneseqn1980s:*
 - 2: geneseqn1990s:*
 - 3: geneseqn2000s:*
 - 4: geneseqn2001as:*
 - 5: geneseqn2001bs:*
 - 6: geneseqn2002as:*
 - 7: geneseqn2002bs:*
 - 8: geneseqn2003as:*
 - 9: geneseqn2003bs:*
 - 10: geneseqn2003cs:*
 - 11: geneseqn2003ds:*
 - 12: geneseqn2004as:*
 - 13: geneseqn2004bs:*
 - 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1197	100.0	1197	4	AAF84670	Aaf84670 Nucleotid
2	1004	83.9	1465	6	AAL47329	Aal47329 Human utr
3	467	39.0	752	9	ADB82145	Adb82145 Human CDN
4	467	39.0	823	6	ABQ89202	Abq89202 Human pro
5	186.6	15.6	36000	12	ADJ84208	Adj84208 Human JNK
6	182	15.2	47804	10	ADC86176	Adc86176 Human GPC
7	180.4	15.1	1145	4	AAF84671	Aaf84671 Nucleotid
8	174.8	14.6	1500	4	AAF84672	Aaf84672 Nucleotid
9	174.8	14.6	110000	8	ABQ83210_0	Abq83210 Human tra
10	171.2	14.3	110000	10	AAL52246_2	Continuation (3 of
11	170.6	14.3	349981	10	ADC87619	Adc87619 Human GPC
12	169.8	14.2	2551	4	AAH18701	Aah18701 Human CDN
13	169	14.1	509	6	ABN61707	Abn61707 Human can
14	168.6	14.1	183178	10	ADL13873	Adl13873 Osteoarth
15	167.8	14.0	133100	12	ADP45594	Adp45594 Human NUM
16	167.8	14.0	191350	14	ADX98571	Adx98571 Human NUM
17	166.8	13.9	549	6	ABN62688	Abn62688 Human can
18	166.8	13.9	612	14	ACL60100	ACL60100 Human col
19	166.4	13.9	23544	4	AAL05829	Aal05829 Human rep

C	20	166.4	13.9	23544	4	ABL98393	AbL98393 Human tes
C	21	165.8	13.9	6637	4	AAI98097	Aai98097 Human neu
C	22	165.6	13.8	31934	4	AAK82215	Aak82215 Human imm
C	23	165	13.8	110000	14	ADZ42285_2	Continuation (3 of
C	24	164.8	13.8	51469	4	AAK78813	Aak78813 Human imm
C	25	164.8	13.8	51469	4	AAK70270	Aak70270 Human imm
C	26	164.8	13.8	51469	4	AAK69322	Aak69322 Human imm
C	27	164.2	13.7	1001	12	ADO35129	Ado35129 Human KCh
C	28	164	13.7	1182	2	AAZ27261	Aaz27261 Human sec
C	29	164	13.7	325791	4	AAS43104	Aas43104 Human Oes
C	30	163	13.6	26277	4	AAK70428	Aak70428 Human imm
C	31	163	13.6	82256	12	ADO34927_3	Continuation (4 of
C	32	162	13.5	113000	9	ABT44365	Abt44365 Partial g
C	33	161.8	13.5	167932	10	ADL13501	Adl13501 Osteoarth
C	34	161.8	13.5	227968	6	ABK83497	Abk83497 Human CDN
C	35	161.8	13.5	227968	12	ADQ18538	Adq18538 Human sof
C	36	160.6	13.4	3164	12	ADQ64115	Adq64115 Novel hum
C	37	160.4	13.4	2132	4	ABA07309	Abao7309 Human pan
C	38	160.4	13.4	2132	4	AAK90468	Aak90468 Human dig
C	39	160.4	13.4	189013	8	ACF62741	Acf62741 Cancer ba
C	40	160.4	13.4	189013	8	ADB20856	Adb20856 MRP1 base
C	41	160.4	13.4	189013	10	ADB87945	Adb87945 Human UGT
C	42	160.4	13.4	189013	10	ADB96928	Adb96928 Human MDR
C	43	160.4	13.4	189013	10	ADB92119	Adb92119 Human MDR
C	44	158.6	13.2	153170	12	ADQ17382	Adq17382 Human sof
C	45	157.8	13.2	106315	11	ACN43966	Acn43966 Human gen

ALIGNMENTS

RESULT 1	AAF84670	standard; DNA; 1197 BP.
ID	AAF84670	
XX	AAF84670;	
AC	29-JUN-2001	(first entry)
DT		
XX		
DE		Nucleotide sequence of utrophin exon 1B and alternative promoter.
XX		
KW		Utrophin; promoter; chromosome 6q24; exon 1B; dystrophin;
KM		muscle-specific transcription; muscular dystrophy; ss.
XX		
OS		Homo sapiens.
XX		
FH		
FT	Key	location/Qualifiers
FT	CDS	1021..1114
FT		/*tag= a
FT		/note= "position 1116 corresponds to a splice to exon 3"
XX		
PN	W0200125461-A1.	
XX		
PD	12-APR-2001.	
XX		
PE	04-OCT-2000; 2000WO-GB003800.	
XX		
PR	04-OCT-1999; 99GB-00023423.	
XX		
PA	(ISIS-) ISIS INNOVATION LTD.	
XX		
PI	Burton E, Tinsley J, Davies K;	
XX		
DR	WPI; 2001-273582/28.	
DR	P-PSDB; AAB67961.	
XX		
PT		Novel nucleic acid comprising promoter for mouse, human utrophin genes,
PT		for promoting tissue-specific transcription of linked sequences and in
PT		screening for substances able to modulate utrophin promoter activity.
XX		
PS	Claim 1; Fig 1; 76pp; English.	
XX		
CC		The present sequence represents exon 1B and an alternative promoter of

CC the utrophin gene. The specification describes human and murine utrophin
CC alternative promoters. Utrophin is a 395 kDa protein encoded by a gene
CC located on chromosome 6q24. The alternative promoter is highly regulated,
CC and has little similarity to the synaptically expressed promoter. The
CC alternative promoter drives transcription of a widely expressed unique
CC first exon that splice into a common full length mRNA at exon 3. This
CC unique exon (called exon 1B) encodes a novel 31 peptide which may be
CC involved in binding to the muscle membrane. Utrophin alternative
CC promoters are useful for promoting tissue-specific, preferably muscle-
CC specific transcription of an operably linked sequence of nucleotides.
CC Utrophin exon 1B polynucleotides and polypeptides are useful in the
CC manufacture of medicament for treating a dystrophin phenotype in a
CC mammal. Up-regulation of utrophin expression may compensate for
CC dystrophin loss in muscular dystrophy patients
XX
SQ Sequence 1197 BP; 383 A; 245 C; 264 G; 305 T; 0 U; 0 Other;

Query Match 100.0%; Score 1197; DB 4; Length 1197;
Best Local Similarity 100.0%; Pred. No. 1e-268;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTCTATTTCACAACAAGCAAGAAAAGAAATGAGAGAGACTAGAAAAGTAGATGTGATC 60
Db 1 TTCTATTTCACAACAAGCAAGAAAAGAAATGAGAGAGACTAGAAAAGTAGATGTGATC 60
QY 61 ATATGAAATATGATTTTCTGCTTTTGGCATGTATGTGTGACACATGCAGAAAGTGAC 120
Db 61 ATATGAAATATGATTTTCTGCTTTTGGCATGTATGTGTGACACATGCAGAAAGTGAC 120
QY 121 AGCAGAGTTGAGAGACCGCTGACCAACATGTTGAAATCCCGTCTCTACTAAACACACA 180
Db 121 AGCAGAGTTGAGAGACCGCTGACCAACATGTTGAAATCCCGTCTCTACTAAACACACA 180
QY 181 CACACACACACACACACACACACACACACACACACACAAATAGCCGGCATGTGTGG 240
Db 181 CACACACACACACACACACACACACACACACACACACAAATAGCCGGCATGTGTGG 240
QY 241 GCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCACAAGAAATGACTTGAACCCAGAGG 300
Db 241 GCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCACAAGAAATGACTTGAACCCAGAGG 300
QY 301 CGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 360
Db 301 CGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 360
QY 361 AAAAATAATGATTAATAAGAGAGCAAGGTGACCAAAAAGAAATAGGCTGAAAAATTT 420
Db 361 AAAAATAATGATTAATAAGAGAGCAAGGTGACCAAAAAGAAATAGGCTGAAAAATTT 420
QY 421 GTCTAAATGGTGGCCTCTTCTTAATAGCTGCATATGTTAAGTTATTTTCCCTAGT 480
Db 421 GTCTAAATGGTGGCCTCTTCTTAATAGCTGCATATGTTAAGTTATTTTCCCTAGT 480
QY 481 AGCGAATTCTAAGGGATGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 540
Db 481 AGCGAATTCTAAGGGATGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 540
QY 541 AACTACTATAGTGAATAATAAGTCAATTTATTTTGAAGTATAGTTAATATGTAACG 600
Db 541 AACTACTATAGTGAATAATAAGTCAATTTATTTTGAAGTATAGTTAATATGTAACG 600
QY 601 AAACCTCTAAGGCCAGTTGTATACCGGCAACGCTTCTAACAATCTTATTTATCTA 660
Db 601 AAACCTCTAAGGCCAGTTGTATACCGGCAACGCTTCTAACAATCTTATTTATCTA 660
QY 661 CGCAGTGGGTAGGGAGGTGGGTGAGTGCCTTCCAGCTGATATCTCAAAAACAGAA 720
Db 661 CGCAGTGGGTAGGGAGGTGGGTGAGTGCCTTCCAGCTGATATCTCAAAAACAGAA 720
QY 721 GCAAAATTATATCTCTGTCAATAGAACATGATAGAGGCCCTTAGTGTGACTATTA 780
Db 721 GCAAAATTATATCTCTGTCAATAGAACATGATAGAGGCCCTTAGTGTGACTATTA 780

QY 781 AAAACAAAAAACTGCTTAAGAGTTTTCAGTACTACAAAGTAACTTCTCTGTGT 840
Db 781 AAAACAAAAAACTGCTTAAGAGTTTTCAGTACTACAAAGTAACTTCTCTGTGT 840
QY 841 GTTTAGAGAGTGGGGTTAGTTAGTCAAGATCCTCTCAATGGGAAAAATAAAGCCACC 900
Db 841 GTTTAGAGAGTGGGGTTAGTTAGTCAAGATCCTCTCAATGGGAAAAATAAAGCCACC 900
QY 901 AAAAAAAAAAAAAAAACCAAAATACACACAGACATCCAGTGTGCAAGTGAAGG 960
Db 901 AAAAAAAAAAAAAAAACCAAAATACACACAGACATCCAGTGTGCAAGTGAAGG 960
QY 961 CTGCTTTTGTGTCCACTTCCCTCCACATCTTTTCCCTCATCTAAGCAGATGAGTG 1020
Db 961 CTGCTTTTGTGTCCACTTCCCTCCACATCTTTTCCCTCATCTAAGCAGATGAGTG 1020
QY 1021 ATGAGCGGCTGGCAGCCACCGTTTCATTGGAAAAAGTGCAGATTGATTGCCAGGG 1080
Db 1021 ATGAGCGGCTGGCAGCCACCGTTTCATTGGAAAAAGTGCAGATTGATTGCCAGGG 1080
QY 1081 CATGTAGCTCTCCAGGCTTGCAGCGATTACAGGTAAGTTGTCAACTGACGACTCC 1140
Db 1081 CATGTAGCTCTCCAGGCTTGCAGCGATTACAGGTAAGTTGTCAACTGACGACTCC 1140
QY 1141 CAGCCAGTGAAGTTTCTTAAGAAAGCTATGAAGACAGGTTCTTCAATCAGTT 1197
Db 1141 CAGCCAGTGAAGTTTCTTAAGAAAGCTATGAAGACAGGTTCTTCAATCAGTT 1197

RESULT 2
AAL47329
ID AAL47329 standard; DNA; 1465 BP.
XX
AC AAL47329;
XX
DT 05-SEP-2002 (first entry)
XX
DE Human utrophin B promoter.

XX
KW Human; utrophin B; promoter; Ets-transcription factor; transactivator;
KW muscular disease; Duchenne muscular dystrophy; Becker muscular dystrophy;
KW gene therapy; muscle wasting; cytostatic; immunostimulant; haemostatic;
KW haemophilia; immune deficiency; cancer; ds.
XX
OS Homo sapiens.

XX
PN W0200236620-A2.
XX
PD 10-MAY-2002.
XX
PF 31-OCT-2001; 2001WO-EP012662.
XX
PR 02-NOV-2000; 2000EP-00123842.
XX
PA (MYOC-) MYOCONTRACT PHARM RES AG.

XX
PI Rueegg MA, Brignuet A;
XX
DR WPI; 2002-500123/53.
XX

PT Novel nucleic acid capable of controlling the expression of a gene and
PT being activated by an Ets-transcription factor related compound, useful
PT for screening candidate compounds capable of regulating transcription.
XX

PS Claim 1; Page 22; 38bp; English.

XX The present invention provides a nucleic acid capable of controlling the
CC expression of a gene and being activated by an Ets-transcription factor
CC (EtsF) related compound. The nucleic acid is useful for screening and/or
CC providing candidate compounds capable of regulating transcription. This
CC involves bringing the nucleic acid into contact with compounds to be
CC screened and detecting the transcriptional activity from the nucleic acid
CC in the presence and absence of the compounds and optionally purifying

Db 74 TTCTATTTCACAACAAGCAAGAAAAAGAAATGAGAGAAGAGCTAGAAAGTAGATGTGATC 133
QY 61 ATATGAATAATGATTTTCTTGTCTTTTGGATGATGTGTGACACATGCAGAGTGAC 120
Db 134 ATATGAATAATGATTTTCTTGTCTTTTGGATGATGTGTGACACATGCAGAGTGAC 193
QY 121 AGCAGAGTTGAGACCAAGCCTGACCAACATGTGAAATCCCGTCTTACTTAAACACACA 180
Db 194 AGCAGAGTTGAGACCAAGCCTGACCAACATGTGAAATCCCGTCTTACTTAAACACACA 253
QY 181 CACACACACACACACACACACACACACACACACACAAATAGCCGGGCGATGGTGTGG 240
Db 254 CACACACACACACACACACACACACACACACACACAAATAGCCGGGCGATGGTGTGG 313
QY 241 GCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGG 300
Db 314 GCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGG 373
QY 301 CGAGGTTGAGTGAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGAGTGA 360
Db 374 CGGNGNTGCAGNGAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGAGTGA 433
QY 361 AAAAATAATGATTAATAAAGAGAGCAAGGTGACCAAAAAGAGATAGGCGTGA 420
Db 434 AAAAATAATGATTAATAAAGAGAGCNGGGTGACCAAAAAGAGATAGNNTGGA 493
QY 421 GTCTAATAGTGG-CCTCTTCTCTTATAGCTGCATATGTGTTAAGTTATTTTCCCTAG 479
Db 494 GTCTAANNNGGGGCCCTCTTCTTATAGCTGCATATGTGNTAAGTTATTTTCCCTA- 552
QY 480 TAGCGAATTTCTAAGGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTGA 539
Db 553 NAGCGANTTCT-AGGATGAANNAATNCTTTTCAGTTTCTTCCCGAGGGNGTATA 611
QY 540 TAACTACTATAGTGAATAATTAAGTCCAAATTATTTCTTGAA 581
Db 612 CCTCTA----TGGGAATAATTAAGCCCAATTTTGTGAGTA 649

RESULT 4
ABQ89202
ID ABQ89202 standard; cDNA; 823 BP.
XX ABQ89202;
AC ABQ89202;
DT 27-SEP-2002 (first entry)
XX
DB Human prostate expressed polynucleotide SEQ ID NO 458.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW ss.
OS Homo sapiens.
XX
PN WO200255700-A2.
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US047349.
XX
PR 07-DEC-2000; 2000US-0254648P.
PR 13-MAR-2001; 2001US-0275688P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Cirkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR WPI; 2002-557824/59.
XX
PT New genes and gene products isolated from human prostate, useful for

PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases.
XX
PS Claim 1; SEQ ID NO 458; 186bp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumor or cancer (e.g. prostate cancer, breast cancer, lung
CC cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC rabbits, horse or human). The polynucleotides and polypeptides are also
CC useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence
XX
SQ Sequence 823 BP; 231 A; 164 C; 185 G; 197 T; 0 U; 46 Other;

Query Match 39.0%; Score 467; DB 6; Length 823;
Best Local Similarity 90.9%; Pred. No. 8.9e-99;
Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;

QY 1 TTCTATTTCACAACAAGCAAGAAAAAGAAATGAGAGAAGAGCTAGAAAGTAGATGTGATC 60
Db 145 TTCTATTTCACAACAAGCAAGAAAAAGAAATGAGAGAAGAGCTAGAAAGTAGATGTGATC 204
QY 61 ATATGAATAATGATTTTCTTGTCTTTTGCATGTATGTGTGACACATGCAGAAAGTGAC 120
Db 205 ATATGAATAATGATTTTCTTGTCTTTTGCATGTATGTGTGACACATGCAGAAAGTGAC 264
QY 121 AGCAGAGTTGAGACCAAGCCTGACCAACATGTGAAATCCCGTCTTACTTAAACACACA 180
Db 265 AGCAGAGTTGAGACCAAGCCTGACCAACATGTGAAATCCCGTCTTACTTAAACACACA 324
QY 181 CACACACACACACACACACACACACACACACACACAAATAGCCGGGCGATGGTGTGG 240
Db 325 CACACACACACACACACACACACACACACACACACAAATAGCCGGGCGATGGTGTGG 384
QY 241 GCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGG 300
Db 385 GCACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAACAAGATGACTTGAACCCAGAGG 444
QY 301 CGAGGTTGAGTGAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGAGTGA 360
Db 445 CGGNGNTGCAGNGAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGAGTGA 504
QY 361 AAAAATAATGATTAATAAAGAGCAAGGTGACCAAAAAGAGATAAGGCTGAAAAATTT 420
Db 505 AAAAATAATGATTAATAAAGAGCNGGGTGACCAAAAAGAGATAAGNTGGAANAATTT 564
QY 421 GTCTAATAGTGG-CCTCTTCTCTTATAGCTGCATATGGTTAAGTTATTTTCCCTAG 479
Db 565 GTCTAANNNGGGGCCCTCTTCTTATAGCTGCATATGNTAAGTTATTTTCCCTA- 623
QY 480 TAGCGAATTTCTAAGGGATGAAGAAGAAATCCTTTTCAGTTTAACTTCCCAAGGTGTGA 539
Db 624 NAGCGANTTCT-AGGATGAANNAATNCTTTTCAGTTTCTTCCCGAGGGNGTATA 682
QY 540 TAACTACTATAGTGAATAATTAAGTCCAAATTATTTCTTGAA 581
Db 683 CCTCTA----TGGGAATAATAAGCCCAATTTTGTGAGTA 720

RESULT 5
ADJ84208/c
ID ADJ84208 standard; DNA; 36000 BP.
XX
AC ADJ84208;
XX

DT 06-MAY-2004 (first entry)
XX
DE Human JNK1 DNA which is a target for antisense therapy.
XX
XX antimicrobial; antiinflammatory; cytostatic; infection; inflammation;
KM tumour formation; antisense therapy; human; JNK1; ds; target.
XX
OS Homo sapiens.
XX
PN WO2004003134-A2.
XX
PD 08-JAN-2004.
XX
PF 12-JUN-2003; 2003WO-US018481.
XX
PR 26-JUN-2002; 2002US-0392020P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Monia BP, Freier SM, Manoharan M, Gaarde WA;
XX
DR WPI; 2004-083026/08.
XX
PT Decreasing levels of a preselected cellular mRNA in a cell, useful for
PT treating or preventing a disease associated with a preselected cellular
PT mRNA, comprises binding to the mRNA an antisense compound targeted to a
PT splice site on the mRNA.
XX
PS Example 17; SEQ ID NO 27; 174bp; English.
XX
CC The invention relates to a novel method for decreasing levels of a
CC preselected cellular mRNA in a cell or tissue comprising binding to the
CC preselected cellular mRNA an antisense compound which is specifically
CC hybridisable with a splice site on the mRNA and which is not a substrate
CC for RNase H when bound to the RNA. The method of the invention has
CC antimicrobial, antiinflammatory and cytostatic applications and may be
CC useful for decreasing levels of a preselected cellular mRNA in a cell or
CC tissue and thus for treating or preventing a disease or condition
CC associated with a preselected cellular mRNA or with a preselected target
CC cellular protein, particularly infection, inflammation or tumour
CC formation. The current sequence is that of the human JNK1 DNA of the
CC invention which is a target for antisense therapy.
XX
SQ Sequence 36000 BP; 10349 A; 6446 C; 6627 G; 12578 T; 0 U; 0 Other;
Query Match 15.6%; Score 186.6; DB 12; Length 36000;
Best Local Similarity 77.9%; Pred. No. 4.8e-33;
Matches 225; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 118 GACACAGAGAGTTCGAGACCACTGACCAATCGCTCTCTACTAAACAC 177
Db 3816 GAGATCAGAGATCGAGACCATCCGGCTAACATGGTGAACCTCTCTACTAAAAAT 3757
QY 178 ACACACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGGTGG 237
Db 3756 ACACACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGGTGG 3697
QY 238 TGGGCACTGTATATCCAGCTACTTGGGAGGCTGAGGCACAGAATGACTTGAACCCAGG 297
Db 3696 CAGGCACCTGTAGCCCCAGCTACTCCGGAGGCTGAGGCGGGAATGGCGTGAACCCGGG 3637
QY 298 AGGCGGAGGTTCAGTGCATGATCATGCCATTGCACTTCAGCTGGGTGACGAGTGA 357
Db 3636 AGGCGGAGGTTCAGTGCATGATCATGCCATTGCACTTCAGCTGGGTGACGAGTGA 3577
QY 358 AAAAAAATATGATATATAAGAGCAGCAGGTGACCAACAAAGAGATA 406
Db 3576 AGACTCTGTCTCAAAAAAACAACAAAAAACAACAAAAAACAACAA 3528

RESULT 6
ADC86176
ID ADC86176 standard; DNA; 47804 BP.

XX
AC ADC86176;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human GPCR gene SEQ ID NO:629.
XX
XX ds; gene; human; GPCR;
KM guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1270724-A2.
XX
PD 02-JAN-2003.
XX
PF 18-JUN-2002; 2002EP-00013517.
XX
PR 18-JUN-2001; 2001JP-00246789.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
DR WPI; 2003-315783/31.
DR P-PSDB; ADC86177.
XX
PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX
PS Claim 1; SEQ ID NO 629; 28bp; English.
XX
CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC invention.
XX
SQ Sequence 47804 BP; 13080 A; 10916 C; 11116 G; 12192 T; 0 U; 500 Other;
Query Match 15.2%; Score 182; DB 10; Length 47804;
Best Local Similarity 82.2%; Pred. No. 6e-32;
Matches 222; Conservative 0; Mismatches 45; Indels 3; Gaps 1;
QY 118 GACACAGAGAGTTCGAGACCACTGACCAACATGCTGAATCCCGTCTCTACTAA--A 174
Db 12662 GAGTTCAGAGTTCGAGACTAGCTGGCCACATAGCAAAACCCATCTCTACTAAACAA 12721
QY 175 CACACACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGG 234
Db 12722 CACACACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGG 12781
QY 235 TGGTGGGCACTGTATATCCAGCTACTTGGGAGGCTGAGGCACAAGAATGACTTGAACCC 294
Db 12782 TGGTGTGCACCTGTATATCCAGCTACTGAGGAGGCTGAGGACAGAGAATCACTTGAACCC 12841
QY 295 AGGAGCGGAGGTTCAGTGCATGATCATGCCATTGCACTTCAGCTGGGTGACGAG 354
Db 12842 GGGAGGTGAGGTTCAGTGCATGATCATGCCATTGCACTTCAGCTGGGTGACGAG 12901
QY 355 TGAATAAATAATGATGATATAAGAGAGC 384
Db 12902 AGCGAAACTCCGTCTCAAAAAAACAACCC 12931

RESULT 7
AAF84671
ID AAF84671 standard; DNA; 1145 BP.

XX AAF84671; (first entry)
XX 29-JUN-2001
DE Nucleotide sequence of utrophin exon 1B and alternative promoter.
XX
KM Utrophin; promoter; chromosome 6q24; exon 1B; dystrophin;
KM muscle-specific transcription; muscular dystrophy; ss.
XX
OS Mus sp.
FH Key Location/Qualifiers
FT CDS 969..1062
FT /*tag= a
FT /note= "the end corresponds to a splice to exon 3"

XX WO200125461-A1.
XX 12-APR-2001.
XX 04-OCT-2000; 2000WO-GB003800.
XX 04-OCT-1999; 99GB-00023423.
XX (ISIS-) ISIS INNOVATION LTD.
XX Burton E, Tinsley J, Davies K;
XX MPI; 2001-273582/28.
XX P-PSDB; AAB67962.

XX Novel nucleic acid comprising promoter for mouse, human utrophin genes,
PT for promoting tissue-specific transcription of linked sequences and in
PT screening for substances able to modulate utrophin promoter activity.
XX
PS Claim 1; Fig 2; 76pp; English.

CC The present sequence represents exon 1B and an alternative promoter of
CC the utrophin gene. The specification describes human and murine utrophin
CC alternative promoters. Utrophin is a 395 kDa protein encoded by a gene
CC located on chromosome 6q24. The alternative promoter is highly regulated,
CC and has little similarity to the synaptically expressed promoter. The
CC alternative promoter drives transcription of a widely expressed unique
CC first exon that splice into a common full length mRNA at exon 3. This
CC unique exon (called exon 1B) encodes a novel 31 peptide which may be
CC involved in binding to the muscle membrane. Utrophin alternative
CC promoters are useful for promoting tissue-specific, preferably muscle-
CC specific transcription of an operably linked sequence of nucleotides.
CC Utrophin exon 1B polynucleotides and polypeptides are useful in the
CC manufacture of medicament for treating a dystrophin phenotype in a
CC mammal. Up-regulation of utrophin expression may compensate for
CC dystrophin loss in muscular dystrophy patients

XX Sequence 1145 BP; 289 A; 236 C; 285 G; 333 T; 0 U; 2 Other;

Query Match 15.1%; Score 180.4; DB 4; Length 1145;
Best Local Similarity 75.5%; Pred. No. 5.3e-32;
Matches 281; Conservative 0; Mismatches 76; Indels 15; Gaps 4;

OY 755 AGAGGCCCTTAGTGTGACTATTAAAAACAACCTGCTTAAGAGTTTCACTGA 814
DB 718 AAAAGCTTAAGCTGTGACTATTAAAAACAAGTACCT--CAAGATTCTTAAGTGA 774
OY 815 CTACAAAGTGAAGTCTCTCTGTGTTAGAGAGGTGGGTTAGTTTACAGATC 874
DB 775 CTGCGGAGTTTAAGTCTCTGTCTG-----AGGGAGGTGAGTTAGATTAGTCAATC 828
OY 875 CTCTCATGGGAAAAATAAAGCCACCAAAAAAATATACACA 934
DB 829 CTCTCGTGGAAAAAATCAAGGAGCTTTAAAAAGAAAAAACAACCAACCTACA 888
OY 935 GGACATCCAGTGTGACGTTGGAAGGCTGTTTGTGTCCACTTCTCCACATCTTTT 994

DB 889 GGACATCCAGTGTGACGTTGCGGGCGGCTTTTG-TGTGATTTCTTCACAGTTT--- 944
OY 995 CCTCATCATCTAAGCAGATGTAGTATGACGGCCTGGCAGCCACCAAGTTTCAATTGA 1054
DB 945 --CCCTCATCTACGCCACTGTAGTATGATGACAGCCTGGCAGCCACCAATTTCGTTGA 1002
OY 1055 AAAAGTGACATGATTTGCTCCAGGGCAGTGTAGCTCTCCAGGCTTGACAGATTACAG 1114
DB 1003 AAAAGTGAGTTGATCTGCTGGGAGGTGCTCTCCAGGCTTGACAGAGATCCCGG 1062
OY 1115 GTAGTTGTCA 1126
DB 1063 GTAGTTGTCA 1074

RESULT 8
ID AAF84672 standard; DNA; 1500 BP.
XX AAF84672;
XX 29-JUN-2001 (first entry)
XX Nucleotide sequence of utrophin exon 1B splice variant.
DE
XX Utrophin; promoter; chromosome 6q24; exon 1B; dystrophin;
KM muscle-specific transcription; muscular dystrophy; ds.
XX
XX Homo sapiens.

FH Key Location/Qualifiers
FT CDS 81..1500
FT /*tag= a
FT /transl_except= (pos: 1500, aa: Asp)

XX WO200125461-A1.
XX 12-APR-2001.
XX 04-OCT-2000; 2000WO-GB003800.
XX 04-OCT-1999; 99GB-00023423.
XX (ISIS-) ISIS INNOVATION LTD.
XX Burton E, Tinsley J, Davies K;
XX MPI; 2001-273582/28.
XX P-PSDB; AAB67963.

PT Novel nucleic acid comprising promoter for mouse, human utrophin genes,
PT for promoting tissue-specific transcription of linked sequences and in
PT screening for substances able to modulate utrophin promoter activity.

XX Disclosure; Fig 8; 76pp; English.

XX The present sequence encodes the utrophin exon 1B splice variant. The
CC specification describes human and murine utrophin alternative promoters.
CC Utrophin is a 395 kDa protein encoded by a gene located on chromosome
CC 6q24. The alternative promoter is highly regulated, and has little
CC similarity to the synaptically expressed promoter. The alternative
CC promoter drives transcription of a widely expressed unique first exon that
CC splice into a common full length mRNA at exon 3. This unique exon (called
CC exon 1B) encodes a novel 31 peptide which may be involved in binding to
CC the muscle membrane. Utrophin alternative promoters are useful for
CC promoting tissue-specific, preferably muscle-specific transcription of an
CC operably linked sequence of nucleotides. Utrophin exon 1B polynucleotides
CC and polypeptides are useful in the manufacture of medicament for treating
CC a dystrophin phenotype in a mammal. Up-regulation of utrophin expression
CC may compensate for dystrophin loss in muscular dystrophy patients

XX Sequence 1500 BP; 451 A; 328 C; 374 G; 347 T; 0 U; 0 Other;

```
Query Match      14.6%; Score 174.8; DB 4; Length 1500;
Best Local Similarity 98.9%; Pred. No. 1.1e-30;
Matches 176; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      941 CCCAGTGTGCAGTTCGAAGGCTGCTTTTGTGTGCCACTTCCTCCACATCTTTTCCAT 1000
Db      1 CCCAGTGTGCAGTTCGAAGGCTGCTTTTGTGTGCCACTTCCTCCACATCTTTTCCAT 60

QY      1001 CATCTAAGCAGATGTAGTGATGAGCGGCTGCGACGCCACCACTTTCATTGGAAGT 1060
Db      61 CATCTAAGCAGATGTAGTGATGAGCGGCTGCGACGCCACCACTTTCATTGGAAGT 120

QY      1061 GCAGATTGATTTGCCAGGCATGTAGCTCTCCAGGCTTGCAAGCGATTACCAAGTAA 1118
Db      121 GCAGATTGATTTGCCAGGCATGTAGCTCTCCAGGCTTGCAAGCGATTACCAAGTGA 178

RESULT 9
ABQ83210_0
WP Sequence split into 4 fragments LOCUS ABQ83210 Accession Abq83210
WP Fragment Name      Begin      End
WP ABQ83210_0          1      110000
WP ABQ83210_1          100001   210000
WP ABQ83210_2          200001   310000
WP ABQ83210_3          300001   397658
ID ABQ83210 Standard; DNA; 397658 BP.
XX
AC ABQ83210;
XX
DT 17-JAN-2003 (first entry)
XX
DE Human transporter protein genomic DNA SEQ ID NO:3.
XX
KW Human; transporter protein; chromosome 2; gene; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(83,T)
FT CDS /standard_name= "single nucleotide polymorphism (SNP)"
FT /tag= a
FT 2192..395291
FT /tag= b
FT /product= "transporter protein"
FT /note= "contains introns"
FT 2192..2267
FT /tag= c
FT /number= 1
FT 2268..3942
FT /tag= d
FT /number= 1
FT /replace(2342,C)
FT /tag= e
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(2620,A)
FT /tag= f
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(2884,A)
FT /tag= g
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(3196,C)
FT /tag= h
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT 3943..4173
FT /tag= i
FT /number= 2
FT 4174..304169
FT /tag= j
FT /number= 2
FT /replace(4531,T)
FT /tag= k
FT /standard_name= "single nucleotide polymorphism (SNP)"
```

```
FT variation replace(5329..5330,A)
FT /tag= l
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(6291,C)
FT /tag= m
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(7001,A)
FT /tag= n
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(7424..7425,A)
FT /tag= o
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(7951,T)
FT /tag= p
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(8053,A)
FT /tag= q
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(8073,G)
FT /tag= r
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(8146,G)
FT /tag= s
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(9134,C)
FT /tag= t
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(9548,G)
FT /tag= u
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(10230,G)
FT /tag= v
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(10777,T)
FT /tag= w
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(10859..10860,A)
FT /tag= x
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(12462,C)
FT /tag= y
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(13062,C)
FT /tag= z
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(13245,C)
FT /tag= aa
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(13334,A)
FT /tag= ab
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(14049,T)
FT /tag= ac
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(14693,G)
FT /tag= ad
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(14695,C)
FT /tag= ae
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(14725,A)
FT /tag= af
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(14777,G)
FT /tag= ag
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(17466,C)
FT /tag= ah
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(18095,T)
FT /tag= ai
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /replace(19740,A)
```

FT	/*tag= aj	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(20813,A)	
FT	/*tag= ak	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(22158,A)	
FT	/*tag= al	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(22227,T)	
FT	/*tag= am	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(22311,G)	
FT	/*tag= an	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(22414,A)	
FT	/*tag= ao	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(23458,G)	
FT	/*tag= ap	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(26694,C)	
FT	/*tag= aq	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(26713,A)	
FT	/*tag= ar	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(27014,A)	
FT	/*tag= as	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(27467,T)	
FT	/*tag= at	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(28916,G)	
FT	/*tag= au	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(29522,T)	
FT	/*tag= av	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(31683,A)	
FT	/*tag= aw	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(33358,T)	
FT	/*tag= ax	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(33387,C)	
FT	/*tag= ay	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(33753,T)	
FT	/*tag= az	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(34779,A)	
FT	/*tag= ba	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(36787. .36788,T)	
FT	/*tag= bb	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(36982,T)	
FT	/*tag= bc	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(40138,T)	
FT	/*tag= bd	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(40309,A)	
FT	/*tag= be	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(40317,T)	
FT	/*tag= bf	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(46648,C)	
FT	/*tag= bg	/standard_name= "single nucleotide polymorphism (SNP) "
FT	replace(49505,G)	
FT	/*tag= bh	

[illegible]

```

RESULT 10
AAL52246_2
Continuation (3 of 4) of AAL52246 from base 200001 (Human genomic DNA for the gene enc
WP Sequence split into 4 fragments LOCUS AAL52246 Accession Aal52246
WP      Fragment Name      Begin      End
WP      AAL52246_0         1          110000
WP      AAL52246_1        100001     210000
WP      AAL52246_2        200001     310000
WP      AAL52246_3        300001     378361

```

[illegible]

```

RESULT 11
ADC87619
ID   ADC87619 standard; DNA; 349981 BP.
XX
XX   AC   ADC87619;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Human GPCR related polynucleotide SEQ ID NO:2072.
XX
KW   ds; human; GPCR; guanosine triphosphate-binding protein coupled receptor;

```


KW	gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	EPI270724-A2.
XX	
PD	02-JAN-2003.
XX	
PF	18-JUN-2002; 2002EP-00013517.
XX	
PR	18-JUN-2001; 2001JP-00246789.
XX	
PA	(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY. (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX	
PI	Suwa M, Asai K, Akiyama Y, Aburatani H;
DR	WPI; 2003-315783/31.
XX	
PT	New polynucleotide, useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor.
PT	
PS	Disclosure; SEQ ID NO 2072; 28bp; English.
XX	
CC	The invention relates to a novel polynucleotide encoding a guanosine triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of the invention may have a use in gene therapy. The polynucleotide and polypeptide are useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor. The protein sequences shown in ADC87618-ADC87623 represent polynucleotide sequences related to the invention.
CC	
SQ	Sequence 349981 BP; 100900 A; 78560 C; 77396 G; 92823 T; 0 U; 302 Other;
XX	
XX	
Query Match	14.3%; Score 170.6; DB 10; Length 349981;
Best Local Similarity	75.6%; Pred. No. 4.6e-29;
Matches 226; Conservative	0; Mismatches 69; Indels 4; Gaps 1;
OY	93 GTATGTGTTGACACATGCAGAAGTGACAGAGAGTTCGAGACCAGCCTGACCAACATG 152
DB	54155 GGAGGCCGAGGCTAGTGGTTCACCTGAGGTTCAGGAATTGAGACCACTGGCCAACATG 54214
OY	153 GTGAATCCCCGTTCTTAATA---ACACACACACACACACACACACACACACACACA 208
DB	54215 GTGAACCCCTGTCTTCACATAAAATAACACACACACACACACACACACACACACACA 54274
OY	209 CACACACACACAATAGCGCGGCATGTGGTGGCACCTGTAATCCAGCTACTTGGAGG 268
DB	54275 CACACACACAATAAGCTGGGCATGTAGCAGGCACCTGTAATCCAGCTACTAAGGAGG 54334
OY	269 CTGAGGCACAGAATGACTTGAACCCAGAGGCGGAGTTGCACTGAGCTGAGATCATGC 328
DB	54335 CTGAGGCACAGAATCGCTTGAATCTGGGAGGCAGAGGTTGCACTGAGCCGAGATCGCAC 54394
OY	329 CATTCACATCCAGCCTGGGTGACGAGTGAAAAAAAATAATGATTAATAAGAGAGCAAG 387
DB	54395 CACTGCACTTTAGCCCTGGGCGACAGAGCAAGACTCCATCTTAAAAAAAAAAAAAAAG 54453
RESULT 12	
ID	AH18701 standard; cDNA; 2551 BP.
XX	
AC	AH18701;
XX	
DT	26-JUN-2001 (first entry)
XX	
DE	Human cDNA sequence SEQ ID NO:18968.
XX	
KX	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS	Homo. sapiens.
XX	
PN	EP1074617-A2.
XX	
PD	07-FEB-2001.
XX	
PF	28-JUL-2000; 2000EP-00116126.
XX	
PR	29-JUL-1999; 99JP-00248036.
PR	27-AUG-1999; 99JP-00300253.
PR	11-JAN-2000; 2000JP-00118776.
PR	02-MAY-2000; 2000JP-00183767.
PR	09-JUN-2000; 2000JP-00241899.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX	
DR	WPI; 2001-318749/34.
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT	length cDNAs defined in the specification, and for the detection and/or
PT	diagnosis of the abnormality of the proteins encoded by the full-length
PT	cDNAs.
XX	
PS	Claim 8; SEQ ID NO 18968; 2537bp + Sequence Listing; English.
XX	
CC	The present invention describes primer sets for synthesising 5602 full-
CC	length cDNAs defined in the specification. Where a primer set comprises:
CC	(a) an oligo-dT primer and an oligonucleotide complementary to the
CC	complementary strand of a polynucleotide which comprises one of the 5602
CC	nucleotide sequences defined in the specification, where the
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC	of an oligonucleotide comprising a sequence complementary to the
CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5'-end sequence/3'-end sequence is selected from those defined in the
CC	specification. The primer sets can be used in antisense therapy and in
CC	gene therapy. The primers are useful for synthesising polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
XX	
SQ	Sequence 2551 BP; 639 A; 635 C; 588 G; 689 T; 0 U; 0 Other;
	Query Match 14.2%; Score 169.8; DB 4; Length 2551;
	Best Local Similarity 75.8%; Pred. No. 1.9e-29;
	Matches 225; Conservative 0; Mismatches 67; Indels 5; Gaps 1;
QY	118 GACAGCAGGAGTTGCAGACCAGCCTGAACCAACTGTGAATCCCGTCTACTTAACAC 177 .
Db	1489 GAGGTCAAGATTTCAGACCAGCCTGACCAACATGATAAACCCTCTCTAATAAAT 1430
QY	178 ACACACACACACACACACACACACACACACACACACAC-----AATAGCCGGCAT 232
Db	1429 ACACACACACACACACACACACACACACACACACACACAAAAAATAGCTGACAT 1370
QY	233 GGTGGTGCGCACCTGTAATCCAGCTACTTGGAGGCTGAGGCCAACAGAATGACTTGAAC 292
Db	1369 GTGGCATGCACCTGTAATCCAGCTAATCCAGCTAATCGAGGCTGAGGCAAGAGAAATCGTTGATC 1310
QY	293 CCAGGAGGCGGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACCTCGGGTGACG 352
Db	1309 CTGGGAGGTGAAGGTTGCAGTGAGCCGAGACCATGCCCCTGCACTCCAGCTGGGAGACA 1250

[illegible]

RESULT 15
ADP45594
ID ADP45594 standard; DNA; 133100 BP.

AC ADP45594;

DT 26-AUG-2004 (first entry)

Human NUMA1/FLJ20625/LOC220074 region gDNA

xx breast cancer; cytostatic; gene therapy; human; ds; gene; SNP;
kw

KW single nucleotide p
KW chromosome 11q13.3.

aa
os Homo sapiens.

Key	Location/Qualifiers
-----	---------------------

FI Valuation
FI
FT
/ *tag= a
" (SNB) "

FT	variation	815
FI		

```

11      /standard_name= "Single nucleotide polymorphism (SNP)"
12      .
13      .
14      .
15      .
16      .
17      .
18      .
19      .
20      .
21      .
22      .
23      .
24      .
25      .
26      .
27      .
28      .
29      .
30      .
31      .
32      .
33      .
34      .
35      .
36      .
37      .
38      .
39      .
40      .
41      .
42      .
43      .
44      .
45      .
46      .
47      .
48      .
49      .
50      .
51      .
52      .
53      .
54      .
55      .
56      .
57      .
58      .
59      .
60      .
61      .
62      .
63      .
64      .
65      .
66      .
67      .
68      .
69      .
70      .
71      .
72      .
73      .
74      .
75      .
76      .
77      .
78      .
79      .
80      .
81      .
82      .
83      .
84      .
85      .
86      .
87      .
88      .
89      .
90      .
91      .
92      .
93      .
94      .
95      .
96      .
97      .
98      .
99      .
100     .
101     .
102     .
103     .
104     .
105     .
106     .
107     .
108     .
109     .
110     .
111     .
112     .
113     .
114     .
115     .
116     .
117     .
118     .
119     .
120     .
121     .
122     .
123     .
124     .
125     .
126     .
127     .
128     .
129     .
130     .
131     .
132     .
133     .
134     .
135     .
136     .
137     .
138     .
139     .
140     .
141     .
142     .
143     .
144     .
145     .
146     .
147     .
148     .
149     .
150     .
151     .
152     .
153     .
154     .
155     .
156     .
157     .
158     .
159     .
160     .
161     .
162     .
163     .
164     .
165     .
166     .
167     .
168     .
169     .
170     .
171     .
172     .
173     .
174     .
175     .
176     .
177     .
178     .
179     .
180     .
181     .
182     .
183     .
184     .
185     .
186     .
187     .
188     .
189     .
190     .
191     .
192     .
193     .
194     .
195     .
196     .
197     .
198     .
199     .
200     .
201     .
202     .
203     .
204     .
205     .
206     .
207     .
208     .
209     .
210     .
211     .
212     .
213     .
214     .
215     .
216     .
217     .
218     .
219     .
220     .
221     .
222     .
223     .
224     .
225     .
226     .
227     .
228     .
229     .
230     .
231     .
232     .
233     .
234     .
235     .
236     .
237     .
238     .
239     .
240     .
241     .
242     .
243     .
244     .
245     .
246     .
247     .
248     .
249     .
250     .
251     .
252     .
253     .
254     .
255     .
256     .
257     .
258     .
259     .
260     .
261     .
262     .
263     .
264     .
265     .
266     .
267     .
268     .
269     .
270     .
271     .
272     .
273     .
274     .
275     .
276     .
277     .
278     .
279     .
280     .
281     .
282     .
283     .
284     .
285     .
286     .
287     .
288     .
289     .
290     .
291     .
292     .
293     .
294     .
295     .
296     .
297     .
298     .
299     .
300     .
301     .
302     .
303     .
304     .
305     .
306     .
307     .
308     .
309     .
310     .
311     .
312     .
313     .
314     .
315     .
316     .
317     .
318     .
319     .
320     .
321     .
322     .
323     .
324     .
325     .
326     .
327     .
328     .
329     .
330     .
331     .
332     .
333     .
334     .
335     .
336     .
337     .
338     .
339     .
340     .
341     .
342     .
343     .
344     .
345     .
346     .
347     .
348     .
349     .
350     .
351     .
352     .
353     .
354     .
355     .
356     .
357     .
358     .
359     .
360     .
361     .
362     .
363     .
364     .
365     .
366     .
367     .
368     .
369     .
370     .
371     .
372     .
373     .
374     .
375     .
376     .
377     .
378     .
379     .
380     .
381     .
382     .
383     .
384     .
385     .
386     .
387     .
388     .
389     .
390     .
391     .
392     .
393     .
394     .
395     .
396     .
397     .
398     .
399     .
400     .
401     .
402     .
403     .
404     .
405     .
406     .
407     .
408     .
409     .
410     .
411     .
412     .
413     .
414     .
415     .
416     .
417     .
418     .
419     .
420     .
421     .
422     .
423     .
424     .
425     .
426     .
427     .
428     .
429     .
430     .
431     .
432     .
433     .
434     .
435     .
436     .
437     .
438     .
439     .
440     .
441     .
442     .
443     .
444     .
445     .
446     .
447     .
448     .
449     .
450     .
451     .
452     .
453     .
454     .
455     .
456     .
457     .
458     .
459     .
460     .
461     .
462     .
463     .
464     .
465     .
466     .
467     .
468     .
469     .
470     .
471     .
472     .
473     .
474     .
475     .
476     .
477     .
478     .
479     .
480     .
481     .
482     .
483     .
484     .
485     .
486     .
487     .
488     .
489     .
490     .
491     .
492     .
493     .
494     .
495     .
496     .
497     .
498     .
499     .
500     .
501     .
502     .
503     .
504     .
505     .
506     .
507     .
508     .
509     .
510     .
511     .
512     .
513     .
514     .
515     .
516     .
517     .
518     .
519     .
520     .
521     .
522     .
523     .
524     .
525     .
526     .
527     .
528     .
529     .
530     .
531     .
532     .
533     .
534     .
535     .
536     .
537     .
538     .
539     .
540     .
541     .
542     .
543     .
544     .
545     .
546     .
547     .
548     .
549     .
550     .
551     .
552     .
553     .
554     .
555     .
556     .
557     .
558     .
559     .
560     .
561     .
562     .
563     .
564     .
565     .
566     .
567     .
568     .
569     .
570     .
571     .
572     .
573     .
574     .
575     .
576     .
577     .
578     .
579     .
580     .
581     .
582     .
583     .
584     .
585     .
586     .
587     .
588     .
589     .
590     .
591     .
592     .
593     .
594     .
595     .
596     .
597     .
598     .
599     .
600     .
601     .
602     .
603     .
604     .
605     .
606     .
607     .
608     .
609     .
610     .
611     .
612     .
613     .
614     .
615     .
616     .
617     .
618     .
619     .
620     .
621     .
622     .
623     .
624     .
625     .
626     .
627     .
628     .
629     .
630     .
631     .
632     .
633     .
634     .
635     .
636     .
637     .
638     .
639     .
640     .
641     .
642     .
643     .
644     .
645     .
646     .
647     .
648     .
649     .
650     .
651     .
652     .
653     .
654     .
655     .
656     .
657     .
658     .
659     .
660     .
661     .
662     .
663     .
664     .
665     .
666     .
667     .
668     .
669     .
670     .
671     .
672     .
673     .
674     .
675     .
676     .
677     .
678     .
679     .
680     .
681     .
682     .
683     .
684     .
685     .
686     .
687     .
688     .
689     .
690     .
691     .
692     .
693     .
694     .
695     .
696     .
697     .
698     .
699     .
700     .
701     .
702     .
703     .
704     .
705     .
```

FT / *tag= c "si:na) a nusi co:ti do no] wmo:hi sm (SNP) "

FT variation.	9715	-
---------------	------	---

```

standard_name= "Single nucleotide polymorphism (SNP)"

```

```
/*tag= e .  
/...[c]e nvc]ectide nc]ymorphism (SNP)'
```

FT	variation	15912	15912
		7452	7452

FT	standard_name=	Single nucleotide polymorphisms (SNP)
10074		

/*tag= g
/tag= g
single nucleotide polymorphism (SNP)
/tag= g

FT variation	19850	19851	19852	19853	19854	19855	19856	19857	19858	19859	19860	19861	19862	19863	19864	19865	19866	19867	19868	19869	19870	19871	19872	19873	19874	19875	19876	19877	19878	19879	19880	19881	19882	19883	19884	19885	19886	19887	19888	19889	19890	19891	19892	19893	19894	19895	19896	19897	19898	19899
FT variation	19850	19851	19852	19853	19854	19855	19856	19857	19858	19859	19860	19861	19862	19863	19864	19865	19866	19867	19868	19869	19870	19871	19872	19873	19874	19875	19876	19877	19878	19879	19880	19881	19882	19883	19884	19885	19886	19887	19888	19889	19890	19891	19892	19893	19894	19895	19896	19897	19898	19899

```

FT      /standard_name= "Single nucleotide polymorphism (SNP)"
      20171

```

```

/*tag= 1
/standard name= "Single nucleotide polymorphism (SNP)

```

FT	variation	20500
----	-----------	-------

```

FT      /standard_name= "Single nucleotide polymorphism (SNP)"
      30536

```

```

/*tag= k
/standard name= "single nucleotide polymorphism (SNP)

```

FT	variation	23187	.
FT		/#tag-]

EM	FT	variation
25289	/standard_name=	single nucleotide polymorphism

```

FT      /*tag= m
FT      /standard name= "Single nucleotide polymorphism (SNP)

```

FT	variation	25470	/*tag= n
FT			

FT	variation	FT	/standard_name-	single	number	FT
FT	28720					

```

FT /-lag= 0
FT /standard name= "Single nucleotide polymorphism (SNP)"

```

FT	variation	29566
E1		

FT	/tag= p
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	30155
FT	/tag= q
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	30752
FT	/tag= r
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	32710
FT	/tag= s
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	32954
FT	/tag= t
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	33725
FT	/tag= u
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	33842
FT	/tag= v
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	36345
FT	/tag= w
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	38115
FT	/tag= x
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	39150
FT	/tag= y
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	40840
FT	/tag= z
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	41969
FT	/tag= aa
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	42045
FT	/tag= ab
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	43785
FT	/tag= ac
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	44444
FT	/tag= ad
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	44579
FT	/tag= ae
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	45386
FT	/tag= af
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	46827
FT	/tag= ag
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	47320
FT	/tag= ah
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	47625
FT	/tag= ai
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	47837
FT	/tag= aj
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	47866
FT	/tag= ak
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	49002
FT	/tag= al
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	49566
FT	/tag= am
FT	/standard_name= "Single nucleotide polymorphism (SNP) "
FT	52058
FT	/tag= an

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 20:24:08 ; Search time 5428 Seconds
(without alignments)
10317.638 Million cell updates/sec

Title: US-10-089-928-1
Perfect score: 1197
Sequence: 1 ttctattcacaacaagca.....cagggtcttctcattcagtt 1197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 segs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199.8	16.7	768	3 BP433396	BP433396 BP433396
2	199.8	16.7	815	5 BW971277	BM971277 BW971277
3	186.2	15.6	671	10 AG168614	AG168614 Pan trogl
4	177.4	14.8	410	8 R63239	R63239 yi07g07.s1
5	176.2	14.7	415	1 AI285383	AI285383 qt64c11.x
6	168.8	14.1	268	2 BF926112	BF926112 CM3-NT017
7	165.8	13.9	651	10 AG000280	AG000280 Homo sapi
8	165	13.8	666	10 AG000280	AG000280 Homo sapi
9	159.8	13.4	1717	4 CNSLTI15Q	BX247957 human ful
10	159.4	13.3	722	5 BQ435695	BQ435695 AGENCOURT
11	157.8	13.2	728	10 AG143162	AG143162 Pan trogl
12	156.2	13.0	554	9 AQ055260	AQ055260 CIT-HSP-2
13	154.2	12.9	464	9 AQ040481	AQ040481 HS_5065_B
14	153.8	12.8	536	5 BU928244	BU928244 AGENCOURT
15	153.6	12.8	593	3 BI584087	BI584087 RH22861.5
16	153.6	12.8	611	3 BI566779	BI566779 RH36522.5
17	153.4	12.8	515	3 BP421422	BP421422 BP421422
18	153.2	12.8	523	9 AQ183529	AQ183529 HS_3197_B
19	153.2	12.8	663	10 AG067676	AG067676 Pan trogl
20	152.8	12.8	625	1 AA594742	AA594742 no03g09.s
21	152.6	12.7	542	3 BI563920	BI563920 RH34666.5
22	151.4	12.6	335	1 AA594694	AA594694 no01g01.s

C	23	150.8	12.6	333	1	AA632711	AA632711 np83c06.s
C	24	150.4	12.6	707	5	BX509705	BX509705 DKF2p686A
C	25	149.4	12.5	231	8	H91330	H91330 yu96h03.s1
C	26	148.4	12.4	523	9	AQ211917	AQ211917 HS_3241_B
C	27	147.6	12.3	354	5	BU945535	BU945535 AGENCOURT
C	28	147.4	12.3	414	1	AA350028	AA350028 EST57300
C	29	147.2	12.3	773	6	CD101811	CD101811 AGENCOURT
C	30	146	12.2	891	5	BX457806	BX457806 BX457806
C	31	146	12.2	895	2	BG618172	BG618172 602645089
C	32	145.4	12.1	656	7	CN273598	CN273598 170006000
C	33	144.6	12.1	420	1	AI717989	AI717989 as93a02.x
C	34	144.4	12.1	647	5	BX954953	BX954953 DKF2p781L
C	35	143.8	12.0	777	1	AV762713	AV762713 AV762713
C	36	142	11.9	399	9	AQ058632	AQ058632 CIT-HSP-2
C	37	142	11.9	467	2	BE143862	BE143862 MRO-HT016
C	38	142	11.9	621	5	BX645545	BX645545 DKF2p781N
C	39	141.8	11.8	578	9	AQ699181	AQ699181 HS_5565_A
C	40	141.8	11.8	852	9	AQ739423	AQ739423 HS_5387_B
C	41	141.4	11.8	667	10	AG090817	AG090817 Pan trogl
C	42	141.4	11.8	5642	4	CR749204	CR749204 Homo sapi
C	43	141.2	11.8	452	1	AA493729	AA493729 nh05g05.s
C	44	140.8	11.8	588	6	CD691266	CD691266 EST7789.h
C	45	140.6	11.7	413	1	AA169830	AA169830 zo92a04.s

ALIGNMENTS

RESULT 1
BP433396
LOCUS
DEFINITION BP433396 full-length enriched swine cDNA library, adult lung Sus
scrofa cDNA clone LNC010094A05 5', mRNA sequence.
ACCESSION BP433396 GI:40423463
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE
AUTHORS 1 (bases 1 to 768)
Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,
Okumura, N., Hamasima, N. and Awata, T.
TITLE PEDF (Pig EST Data Explorer): construction of a database for ESTs
derived from porcine full-length cDNA libraries
JOURNAL Nucleic Acids Res. 32 (1), D484-D488 (2004)
PUBMED
COMMENT
Contact: Hirohide Uenishi
Animal Genome Laboratory, Genome Research Department
National Institute of Agrobiological Sciences
2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
Tel: +81-29-838-8627
Fax: +81-29-838-8627
Email: huenishi@affrc.go.jp
EST project with full-length enriched cDNA libraries carried out in
Animal Genome Research Program (Japan) by National Institute of
Agrobiological Sciences and STAF-Institute
Single pass sequencing of clones derived from oligo-capped cDNA
library
Vector sequences were eliminated by RepeatMasker version 2002/07/13
and crossmatch version 0.990319
Low quality bases were trimmed based on the quality values.

FEATURES

source
1..768
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="LNC010094A05"
/tissue_type="lung"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult
lung"

Email: cgapbs-remail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

Query Match 13.4%; Score 159.8; DB 4; Length 1717;
Best Local Similarity 84.1%; Pred. No. 1.2e-18;
Matches 196; Conservative 0; Mismatches 27; Indels 10; Gaps 1;

QY 117 TGACAGCAGAGTTGAGACCAAGCTGACCAACATGTTGAAATCCCGTCTACTTAA--- 173
Db 1467 TGAGCTCAGGAGTTGAGCCAGCCTGGCCACATGGCGAACCCTGCTCTATTAAAA 1526

QY 174 -----ACACACACACACACACACACACACACACACACACACACACAAATAGCC 226
Db 1527 TACACACACACACACACACACACACACACACACACACACACACAAATTAGCT 1586

QY 227 GGGCATGTGTGGGCACTGTATCCAGCTACTTGGAGGCTGAGGCACAAGATGAC 286
Db 1587 GGGCTGTGTGGCGGACCGCTGTATCCAGCTACTTGAAGGCTGAGGCAGAGATCAC 1646

QY 287 TTGAACCCAGAGGCGGAGTTGACGTGAGATCATGCCATTGCACCTCC 339
Db 1647 TTGAACCCGAGGAGTTGAGGTTGACGTAGCCGAGATCACGCCACTGTACTTC 1699

RESULT 10
LOCUS BQ435695 722 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT 7838732 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6101634
5', mRNA sequence.

ACCESSION BQ435695
VERSION BQ435695.1 GI:21174771
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 722)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2334 row: P column: 19
High quality sequence stop: 571.

FEATURES
source
1..722
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6101634"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGGCCGCGGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."

ORIGIN
Query Match 13.3%; Score 159.4; DB 5; Length 722;
Best Local Similarity 69.8%; Pred. No. 1.6e-18;
Matches 231; Conservative 0; Mismatches 96; Indels 4; Gaps 1;

QY 21 AGAAAAAGATGAGAGAGACTAGAAAGTAGATGATCATATGAATAATGATTTCTCT 80
Db 32 AGAAAAAGAAATACAAATATGAACACAGAGTTGGCCAAATACATGCTCATGCTGTAAAT 91

QY 81 TGCTTTTGCATGTATGTGTGGACACATGCAGAGTGAACAGCAGAGTTGCAGCCAGC 140
Db 92 CCCAGCACTTTGGAGGCGCAAGGCTGTGTGATCATCTGAGGTCAGGATTTTGAGCCAGC 151

QY 141 CTGACCAACATGTGTAAATCCCGTCTCTACTTAAACACACACACACACACACACACA 200
Db 152 CTGGCCACATGTGTAAACCCCATCTCTACTTAAATAACACACACACACACACACAC 207

QY 201 CACACACACACACACACACATAGCCGGGCATGTGTGGCACCCTGTATCCAGCTAC 260
Db 208 CACACACACACACACACAAATTAGCCAGGCATGTGTGGCAGCTGTAGTCCCACTCC 267

QY 261 TTGGAGGCTGAGGCAAGAAATGACTTGAACCCAGAGCGGAGGTTGCAGTGAGCTGA 320
Db 268 TCAGGAGGCGGAGGACAGGAATGCTTGAACCTAGAGAGTGAAGGCTGACGTGAGCCGA 327

QY 321 GATCATGCCATTGCACTCCAGCCTGGGTGAC 351
Db 328 GATCATGTCAGTGCACTCCAGCCTAAGCAAC 358

RESULT 11
LOCUS AG143162 728 bp DNA linear GSS 08-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-003M09.TJ, genomic survey
sequence.

ACCESSION AG143162
VERSION AG143162.1 GI:16672840
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library RP43-43
Unpublished
2 (bases 1 to 728)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpses@gsc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library RP43-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
location/Qualifiers
1..728
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-003M09.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP43-43 Chimpanzee Male BAC library"

ORIGIN
Query Match 13.2%; Score 157.8; DB 10; Length 728;

Best Local Similarity 83.7%; Pred. No. 3.1e-18;
Matches 195; Conservative 0; Mismatches 27; Indels 11; Gaps 1;

QY 118 GACAGCAGAGTTCGAGACCAAGCTGACCAACATGCTGAATCCGCTCTACTAAACAC 177
|||
Db 143 GAGGTCAAGAGTTCAGATCAGCTGGCCACATGTGTAACCCCGCTCTACTAAA--- 199

QY 178 ACACACACACACACACACACACACACACACACACACATAGCCGGGATGTGG 237
|||
Db 200 -----ACACACACACACACACACACACACACACACACACATTAAGCAGGATGTGG 251

QY 238 TGGGCACCTGTATCCAGCTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAGG 297
|||
Db 252 CGACACCTGTAGTCCAGCTACTCGGAGGCTAAGGTAGAGAATCGGTTGAACCCAGG 311

QY 298 AGCGGAGGTTGAGTGAAGTGAATCATGCTTGCATTCAGCTCGGCTGGGTGA 350
|||
Db 312 AGGCAGAGGTTGACAGTGAAGTGAATCGCACCACTGCACTCCAGCTGGGCGA 364

RESULT 12
AQ055260 554 bp DNA linear GSS 30-JUL-1998
LOCUS
DEFINITION CIT-HSP-234106.TF CIT-HSP Homo sapiens genomic clone 234106,
genomic survey sequence.

ACCESSION AQ055260
VERSION AQ055260.1 GI:3351866
KEYWORDS GSS.
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

TITLE
JOURNAL
COMMENT
1 (bases 1 to 554)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Other_GSSs: CIT-HSP-234106.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..554
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="234106"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN

Query Match 13.0%; Score 156.2; DB 9; Length 554;
Best Local Similarity 83.0%; Pred. No. 6.3e-18;
Matches 190; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 123 CAGAGTTGAGACCAAGCTGACCAACATGCTGAATCCGCTCTACTAAACACACACA 182
|||
Db 272 CAGAGATCAAGACCATCCTGCTAACACGCTGAAGCCCATCTTCTACTAAACACACA 331

QY 183 CACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGGCC 242
|||
Db 332 CACACACACACACACACACACACACACACACACACA-ATTAGCCAGGTGTGTGGCACGC 390

QY 243 ACCTGTAATCCAGCTACTTGGAGGCTGAGGCACAAAGATGACTTGAACCCAGAGCG 302
|||
Db 391 ACCTGTAGACCCAGCTGCTTGGAGGCTGAGGCAGAGAATTAATCGAACCAGGAAGCA 450

QY 303 GAGGTGACGTGAGCTGAGATCATGCTTGCATTCAGCTCGGCTGGGTGAC 351
|||
Db 451 GAGGTGACGTGAGCCAGATCAAGCCACCGCATCTCGGCTGGGCGAC 499

RESULT 13
AQ404481 464 bp DNA linear GSS 13-MAR-1999
LOCUS
DEFINITION HS_5069_B2_C08_T7A RPCI-11 Human Male BAC library Homo sapiens
genomic clone Plate=645 Col=16 Row=F, genomic survey sequence.

ACCESSION AQ404481
VERSION AQ404481.1 GI:4415261
KEYWORDS GSS.
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

TITLE
JOURNAL
COMMENT
1 (bases 1 to 464)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end web Server:
http://www.htsc.washington.edu
Plate: 645 row: F column: 16.
Seq primer: T7
Class: BAC ends
High quality sequence stop: 464.

FEATURES
source
Location/Qualifiers
1..464
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=645 Col=16 Row=F"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN

Query Match 12.9%; Score 154.2; DB 9; Length 464;
Best Local Similarity 69.5%; Pred. No. 1.5e-17;
Matches 242; Conservative 0; Mismatches 94; Indels 12; Gaps 2;

QY 117 TGACAGCAGAGTTCGAGACCAAGCTGACCAACATGCTGAATCCGCTCTACTAACA 176
|||

[illegible]

RESULT 14
BU928244
LOCUS
DEFINITION
BU928244 536 bp mRNA linear EST 18-OCT-2002
AGENCOURT_10435920 NIH_MGC_126 Homo sapiens CDNA clone
IMAGE:6653969 5', mRNA sequence.

ACCESSION	BU928244	GI:24116974
VERSION	BU928244.1	
KEYWORDS	EST.	

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 536)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLCM2904 row: n column: 17
High quality sequence stop: 434.

FEATURES	Location/Qualifiers
source	1. .536

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6653969"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_1ib="NIH MGC 126"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
Site_2: SfiI (ggccgctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGACAGAGTGCCATTACGCGCG-3' and
5'-ATCTAGAGGCCGAGCGCGCCGACATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5-1

```

kb size fraction (other fractions present in NIH_MGC_127 and NIH_MGC_128) . Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC Library. "

Query Match	12.8%	Score 153.8;	DB 5;	Length 536;
Best Local Similarity	82.9%;	Pred. No. 1.7e-17;		
Matches 203; Conservative	0;	Mismatches 32;	Indels 10;	Gaps 2;

[illegible]

OY 174 --ACACAACACACACACACACACACACACACAC-----AATTAGCC 226
| | | | | | | | | | | | | | | | | | | | | | | |
DB 296 ACACAACACACACACACACACACACACACATATAAAATTGAC 355

OY 227 GGGCATGCTGTGGCCACCTGTATCCCACTTCTGGAGGCTGAGGCCAACAAGATGAC 286
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 AGGTATGCTGGCGGCCTATAATTCCAATTTCAGGAGGCTGAGGCAGAGAATTGC 415

[illegible]

Oy	347	GTGAC	351
Db	476	GCAAC	480

RESULT 15					
B1584087					
LOCUS					
DEFINITION					
	B1584087	593 bp	mrna	linear	EST 06-SEP-2001
	RH22861.5	prime RH	Drosophila	melanogaster	normalized Head pflc-1
	Drosophila	melanogaster	cdna	clone RH22861 5,	mrna sequence.

ACCESSION	BI584087	
VERSION	BI584087.1	GI:15475509
KEYWORDS	EST.	
SOURCE	Drosophila melanogaster	(fruit fly)
ORGANISM	Drosophila melanogaster	

REFERENCE
AUTHORS

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 593)

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,

TITLE
BDGP/HHMI RH Drosophila EST Project
JOURNAL
Unpublished (2001)
COMMENT
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Plate: RH.228 row: F column: 1
High quality sequence stop: 540.

FEATURES
location/Qualifiers
1. .593

```

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH22861"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DH5-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pFlc-1"
/note="Organ: head; Vector: pFlc1; Site 1: XhoI; Site 2:

```


Db 99133 AGGCGAGCTTCAGTGAGCCAAAGATCAGCCACTGCACTCTAGCCTGGGCGACAGAGTG 99074

OY 358 AAAAAAAAAATGATTAATAAGAGAGCAAGGTGACCACAAAAAGAGATA 406

Db 99073 AGACTCTGTCTCAAAAAAAAACAAAAAAACAATAAAGCAAAA 99025

RESULT 2

```

US-09-949-016-17103
; Sequence 17103, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17103
; LENGTH: 96690
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17103

```

Query Match	15.2%	Score 182;	DB 3;	Length 96690;
Best Local Similarity	78.8%	Pred. No. 2.9e-36;		
Matches 230; Conservative	0;	Mismatches 60;	Indels 2;	Gaps 1;

QY	117	TGACAGCAGGAGTTGAGAGCAGCCTGACCAACATGTTGAATCCCGTCTCTACTAACA	176
Db	73578	TGAGGTCAGGAGTTGAGAGCCAGCCTGGCTAACATGTTGAACCCTATCTTACCACACA	73637
QY	177	CACACACACACACACACACACACACACACACACACAA--TAGCCGGCGCATGG	234
Db	73638	CACACACACACACACACACACACACACACACACACAAATTAGCTGGCGCTGG	73697
QY	235	TGGTGGCACCCTGTATCCCACTTACTTGGGAGGCTGAGCAAGAATGACTTTGAACC	294
Db	73698	TAGTGACACACCTGTAGTCCCACTACTCGGAGGCTGAGGACGAGATCACCTGAACAT	73757
QY	295	AGGAGGCGGAGGTTGCAGTGAAGTCATGATGCGCATTTGCACTCCAGCCTGGGTGACGAG	354
Db	73758	GGGAGGACAGAGGCTGCACTGAGCCAAAGATCATGCCATTGCACTCCAGCCTGGGTGACAGA	73817
QY	355	TGAAAAAAAAAATAATGATAATAAGAGAGCAAGGTGACCACAAAAAGAGATA	406
Db	73818	GTGAGGCTCCATCTCAAAAAAAAAAAAAAAAAAGAAAGAAAAAGAAAAAAA	73869

RESULT 3

US-09-949-016-15497/c
; Sequence 15497, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15497
; LENGTH: 245286
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(245286)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15497

```

Query Match	14.4%;	Score 172.6;	DB 3;	Length 245286;
Best Local Similarity	83.4%;	Pred. No. 1e-33;		
Matches 196; Conservative	0;	Mismatches 39;	Indels 0;	Gaps 0;

OY	117	TGACAGCAGAGATTCCAGACCAGCCTGACCAATCATGTGAATCCCCTCTTACTTAACA	176
Db	235796	TGAGGTGAGAGTTCAAAGATCACGCTGGCCAATGGTGAACCCCATCTCTACTCAAA	235737
OY	177	CACACACACACACACACACACACACACACACACACACACATAAGCCGGCATGTC	236
Db	235736	TACACACACACACACACACACACACACACACACACACACATTAACCGTGCGTC	235677
OY	237	GTCGCACCTGTAATCCAGCTACTTGGAGCGTGAGGCACAAGAATGACTTGAACCCAG	296
Db	235676	GTCGTGCCCTGTAATCCAGTACTTGGGAGCGTGGTTCAAGATATATCATTTGAACCCAG	235617
OY	297	GAGCGGAGGTTGCATGAGCTGAGATCATGCCATTGCACCTCCAGCCTGGGTGAC	351
Db	235616	GAAGCAGAGGTTGTAGTAGAGCCAAGATCATGCCACCGCATGCCAGCCTAGGTGAC	235662

RESULT 4

```

US-09-949-016-133886/c
; Sequence 133886, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133886
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133886

```

Query Match	14.4%;	Score 172.2;	DB 3;	Length 601;
Best Local Similarity	83.0%;	Pred. No. 1.5e-34;		
Matches 195;	Conservative 1;	Mismatches 39;	Indels 0;	Gaps 0;

Qy	117	TGACAGCAGGAGTTCGAGACAGCCTGACCCAACATGGTGAATCCCGTCTCTACTAACA	176
Db	403	TGAGGTCAAGAGTTCAAGATCAGCCTGGCCACATGGTGAATCCCATCTCTACTCAAAA	344
Qy	177	CACACACACACACACACACACACACACACACACACATAGCGCGGCATGCTG	236
Db	343	TACACACACACACACACACACACACACACACACACACATTAACCAAGGTGCGGTG	284
Qy	237	GTGGCACCTGTAACTCCACGTACTTGGGAGGCTGAGGCACAAAGATGAAGCTGAACCCAG	296

Db 283 GTGTGCTGCTTAATCCAGTACTTGGAGGCTGGTTCAGGATATCATTTGAACCCAG 224
QY 297 GAGCGGAGGTTGAGTGAAGCTGAGATCATGCTTGCCTCCAGCCTGGTGAC 351
Db 223 GAAGCAGAGGTTGTAGTGAGCCAAAGATCATGCCACCGCATGCGCTAGTGAC 169

RESULT 5

US-09-949-016-130802/c
; Sequence 130802, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130802
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-130802

Query Match 14.3%; Score 171.6; DB 3; Length 601;
Best Local Similarity 77.8%; Pred. No. 2.1e-34;
Matches 207; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 117 TGACAGCAGAGTTGAGACGACCTGACCAACATGTTGAATCCGTTCTTACTAACA 176
Db 556 TGAGGTGAGGAGTTGAGACCAAGTGTGCCAACATGTGAAACTCGTCTTACTAATAA 497
QY 177 CACACACACACACACACACACACACACACACACACACATAGCCGGCATGTG 236
Db 496 AAAAAAATACACATACACACACACACACATACACACACAATAATTAGTGTG 437
QY 237 GTGGCACCCTGTAATCCAGCTACTTGGAGGCTGGCAGACAGATGACTTGAACCCAG 296
Db 436 GCGTGTGCCGTAAATCCAGCAACCCAGAGGCTGAGGAGAGATCACTGAAACCCGG 377
QY 297 GAGCGGAGGTTGAGTGAAGCTGAGATCATGCTTGCCTCCAGCCTGGTGAAGTG 356
Db 376 GAGGTGAGGCTGAGTGAAGCTGAGATCATGCTTGCCTCCAGCCTGGTGAAGTG 317
QY 357 AAAAAAATATGATTAATAAGAGA 382
Db 316 AAGACTCTGTCTCAAAAAA 291

RESULT 6

US-09-949-016-15427
; Sequence 15427, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15427
; LENGTH: 81585
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(81585)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15427

Query Match 14.3%; Score 171.6; DB 3; Length 81585;
Best Local Similarity 77.8%; Pred. No. 1.2e-33;
Matches 207; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 117 TGACAGCAGAGTTGAGACGACCTGACCAACATGTTGAATCCGTTCTTACTAACA 176
Db 49394 TGAGGTGAGGAGTTGAGACCAAGTGTGCCAACATGTTGAATCCGTTCTTACTAATAA 49453
QY 177 CACACACACACACACACACACACACACACACACACACATAGCCGGCATGTG 236
Db 49454 AAAAAAATACACATACACACACACATACACACACACAATAATTAGTGTG 49513
QY 237 GTGGCACCCTGTAATCCAGCTACTTGGAGGCTGAGGACCAAGAATGACTTGAACCCAG 296
Db 49514 GCGTGTGCCGTAAATCCAGCAACCCAGAGGCTGAGGACCAAGAATCACTGAACCCCG 49573
QY 297 GAGCGGAGGTTGAGTGAAGCTGAGATCATGCTTGCCTCCAGCCTGGTGAAGTG 356
Db 49574 GAGGTGAGGCTGAGTGAAGCTGAGATCATGCTTGCCTCCAGCCTGGTGAAGTG 49633
QY 357 AAAAAAATATGATTAATAAGAGA 382
Db 49634 AAGACTCTGTCTCAAAAAA 49659

RESULT 7

US-09-949-016-179542/c
; Sequence 179542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179542
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-179542

Query Match 14.2%; Score 169.6; DB 3; Length 601;
Best Local Similarity 78.9%; Pred. No. 6.8e-34;
Matches 213; Conservative 1; Mismatches 55; Indels 1; Gaps 1;

QY 117 TGACAGCAGAGTTGAGACGACCTGACCAACATGTTGAATCCGTTCTTACTAACA 176
Db 535 TGAGGTGAGGAGTTCAAGACCAAGCTGGCCCAACATGTTGAATCCGTTCTTACTAATAA 476
QY 177 CACACACACACACACACACACACACACACACACACACATAGCCGGCATGTG 236

Db 475 TACACACACACACACACACACACACACACACACA-AAATCGCCTGGGGCATGTG 417

QY 237 GTGGGCACCTGTAATCCCACTTGGGAGGCTGAGGCCAAGAATGACTTGAACCCAG 296

Db 416 GTGAGCGCCTGTAATCCTAGCTGCTTGGGAGGCTGAGGCCAGGAATTGCTTAAACCTGG 357

QY 297 GAGGCGGAGGTTGCAGTGAGCTGAGATCATGCCATTGCCACTCCAGCCTGGGTGACGAGTG 356

Db 356 GAGGCGGAGGTTGCAGTGAACCGAGATCGCGCCACTGCACTCCAAGCCTGGGCGACRGAGC 297

QY 357 AAAAAAAAAATAATGATAATAAGAGGCCAA 386

Db 296 AAGACTCTGTTAAAAAAGAAAAAGAAAAA 267

```

RESULT 8
US-09-949-016-11935/c
; Sequence 11935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11935
; LENGTH: 57914
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11935

```

[illegible]

RESULT 9
US-09-949-016-16921/c
; Sequence 16921, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

: FILE REFERENCE: CL001307
:
: CURRENT APPLICATION NUMBER: US/09/949,016
:
: CURRENT FILING DATE: 2000-04-14
:
: PRIOR APPLICATION NUMBER: 60/241,755
:
: PRIOR FILING DATE: 2000-10-20
:
: PRIOR APPLICATION NUMBER: 60/237,768
:
: PRIOR FILING DATE: 2000-10-03
:
: PRIOR APPLICATION NUMBER: 60/231,498
:
: PRIOR FILING DATE: 2000-09-08
:
: NUMBER OF SEQ ID NOS: 207012
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 16921
:
: LENGTH: 57936
:
: TYPE: DNA
:
: ORGANISM: Human
:
US-09-949-016-16921

```

[illegible]

```

RESULT 10
US-09-949-016-48701/c
? Sequence 48701, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THERE
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for windows Version 4.0
? SEQ ID NO 48701
? LENGTH: 601
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-48701

```

Query Match	'	14.1%;	Score 168.2;	DB 3;	Length 601;
Best Local Similarity		78.3%;	Pred. No. 1.6e-33;		
Matches 217;	Conservative	1;	Mismatches 49;	Indels 10;	Gaps 1;
QY	138 AGCTGACCAACATGTGTAATCCGTCCTACTAA-----ACACACACACACAC	187			

Db 497 AGCTGGCTAACATGGTGAAACCCCATCTCTACTAAATAACACACACACACACACAC 438
Qy 188 ACACACACACACACACACACACACACACATAGCCGGCATGGTGGGCACCTG 247
Db 437 ACACACACACACACACACACACACACACAAATTAGCCGGCATGGTGGGCACCTG 378
Qy 248 TAATCCAGCTACTTGGAGGCTGAGGCACAGAATGACTTGAACCCAGAGCCGAGGT 307
Db 377 TAATCCAGCTACTCCGAGGCTGAGGCAGAGAATCGCTTGAACCCGAGAGCCGAGGT 318
Qy 308 TGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCTGGGTGAGGAGTGAATAATA 367
Db 317 TGCAGTGAGCCGAGATTGCGCCATTGCACTGCACTGGGTGACAAAGAGCGAGATTGTCT 258
Qy 368 ATGATAATAAAGAGCAAGGTGACCAAAAAGAGAA 404
Db 257 CAAAAAAAAAAAAAAAAAGAAAAAGAAAAA 221

RESULT 11
US-09-949-016-48700/C
; Sequence 48700, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48700
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48700

Query Match 14.0%; Score 167; DB 3; Length 601;
Best Local Similarity 78.3%; Pred. No. 3.2e-33;
Matches 217; Conservative 0; Mismatches 50; Indels 10; Gaps 1;
Qy 138 AGCTGACCAACATGTTGAATCCCGTCTCTACTAA-----ACACACACACACAC 187
Db 377 AGCTGGCTAACATGGTGAAACCCCATCTCTACTAAATAACACACACACACACAC 318
Qy 188 ACACACACACACACACACACACACACACATAGCCGGCATGGTGGGCACCTG 247
Db 317 ACACACACACACACACACACACACACAAATTAGCCGGCATGGTGGGCACCTG 258
Qy 248 TAATCCAGCTACTTGGAGGCTGAGGCACAGAATGACTTGAACCCAGAGCCGAGGT 307
Db 257 TAATCCAGCTACTCCGAGGCTGAGGCAGAGAATCGCTTGAACCCGAGAGCCGAGGT 198
Qy 308 TGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCTGGGTGAGGAGTGAATAATA 367
Db 197 TGCAGTGAGCCGAGATTGCGCCATTGCACTGCACTGGGTGACAAAGAGCGAGATTGTCT 138
Qy 368 ATGATAATAAAGAGCAAGGTGACCAAAAAGAGAA 404
Db 137 CAAAAAAAAAAAAAAAAAGAAAAAGAAAAA 101

RESULT 12
US-09-949-016-13182
; Sequence 13182, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13182
; LENGTH: 37875
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13182

Query Match 14.0%; Score 167; DB 3; Length 37875;
Best Local Similarity 78.3%; Pred. No. 1.4e-32;
Matches 217; Conservative 0; Mismatches 50; Indels 10; Gaps 1;
Qy 138 AGCTGACCAACATGTTGAATCCCGTCTCTACTAA-----ACACACACACACAC 187
Db 10975 AGCTGGCTAACATGGTGAAACCCCATCTCTACTAAATAACACACACACACACAC 11034
Qy 188 ACACACACACACACACACACACACACACATAGCCGGCATGGTGGGCACCTG 247
Db 11035 ACACACACACACACACACACACACACAAATTAGCCGGCATGGTGGGCACCTG 11094
Qy 248 TAATCCAGCTACTTGGAGGCTGAGGCACAGAATGACTTGAACCCAGAGCCGAGGT 307
Db 11095 TAATCCAGCTACTCCGAGGCTGAGGCAGAGAATCGCTTGAACCCGAGAGCCGAGGT 11154
Qy 308 TGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCTGGGTGAGGAGTGAATAATA 367
Db 11155 TGCAGTGAGCCGAGATTGCGCCATTGCACTGCACTGGGTGACAAAGAGCGAGATTGTCT 11214
Qy 368 ATGATAATAAAGAGCAAGGTGACCAAAAAGAGAA 404
Db 11215 CAAAAAAAAAAAAAAAAAGAAAAAGAAAAA 11251

RESULT 13
US-09-768-185A-1
; Sequence 1, Application US/09768185A
; Patent No. 6818758
; GENERAL INFORMATION:
; APPLICANT: Cassel, Michael et al
; TITLE OF INVENTION: Estrogen receptor beta variants and
; TITLE OF INVENTION: methods of detection thereof
; FILE REFERENCE: CL000280
; CURRENT APPLICATION NUMBER: US/09/768,185A
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 09768185
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 325791
; TYPE: DNA
; ORGANISM: HUMAN
US-09-768-185A-1

Query Match 13.7%; Score 164; DB 3; Length 325791;
Best Local Similarity 77.3%; Pred. No. 1.8e-31;
Matches 215; Conservative 0; Mismatches 55; Indels 8; Gaps 1;
Qy 117 TGCAGCAGAGGTTGAGACCAGCCTGACCAACATGGTGAATCCGCTCTACTAA---- 173
Db 117 TGCAGCAGAGGTTGAGACCAGCCTGACCAACATGGTGAATCCGCTCTACTAA---- 173


```
Db 51563 TGAGGTGAGAGTTGAGCCCGAGCTGGCCCAACATGGCGAAACCCGCTCTATTAAAAA 51622
QY 174 -----ACACACACACACACACACACACACACACACACACACATAGCCGG 228
Db 51623 TACACACACACACACACACACACACACACACACACACACACACATAGCTGG 51682
QY 229 GCATGTTGGTGGGACCTGTAATCCAGCTACTTGGAGGCTGAGGCACAGATGACTT 288
Db 51683 GCGTGTGGGGGAGCGCCTGTAATCCAGCTACTTGGAGGCTGAGGCAGAGATGACTT 51742
QY 289 GAACCCAGAGGCGGAGGTTGCACTGAGCTGAGATCATGCCATTGCCACTCCAGCCTGGCT 348
Db 51743 GAACCCGAGAGGTTGAGGTTGCACTGAGCTGAGATCATGCCACTGTACTTCTAGCCTGGG 51802
QY 349 GACGAGTGAATAAATAATGATTAATAAGAGAGCAA 386
Db 51803 TTACAGAGTGAGACTTCATCTCAAAAAA 51840
```

RESULT 14

```
US-09-949-016-12208/c
; Sequence 12208, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12208
; LENGTH: 30244
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12208
```

```
Query Match 13.6%; Score 163; DB 3; Length 30244;
Best Local Similarity 73.4%; Pred. No. 1.4e-31;
Matches 224; Conservative 0; Mismatches 75; Indels 6; Gaps 1;

QY 123 CAGGAGTTGAGACCGCTGACCAACATGTTGAAATCCGCTCTACTAAACACACACA 182
Db 11021 CAGGAGTTCAAGACCGCTGGCCCAACATGTTGAAATCCCATCTCTACTAAAA-----A 10968
QY 183 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTTGGGC 242
Db 10967 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCGGTGGAGGC 10908
QY 243 ACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAAGATGACTTGAACCCAGAGGCG 302
Db 10907 GCCTGTAATCCAGCTACTTGGGAGGCTGAGGCAAGATGACTTGAACCCAGAGGCGCA 10848
QY 303 GAGGTTGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
Db 10847 GAGGTTGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10788
QY 363 AAATATGATTAATAAGAGCAAGGTGACCAAAAGAGATAGGCTGAAAAATTGT 422
Db 10787 CTGTCTCAAAAAAATAATTAATAAGTACACAAAAAATTAAAAATAATAAACACATGT 10728
QY 423 CTAAA 427
Db 10727 AGACA 10723
```

RESULT 15

```
US-09-949-016-13550/c
; Sequence 13550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13550
; LENGTH: 30245
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13550
```

```
Query Match 13.6%; Score 163; DB 3; Length 30245;
Best Local Similarity 73.4%; Pred. No. 1.4e-31;
Matches 224; Conservative 0; Mismatches 75; Indels 6; Gaps 1;

QY 123 CAGGAGTTGAGACCGCTGACCAACATGTTGAAATCCGCTCTACTAAACACACACA 182
Db 11021 CAGGAGTTCAAGACCGCTGGCCCAACATGTTGAAATCCCATCTCTACTAAAA-----A 10968
QY 183 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTTGGGC 242
Db 10967 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCGGTGGAGGC 10908
QY 243 ACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAAGATGACTTGAACCCAGAGGCG 302
Db 10907 GCCTGTAATCCAGCTACTTGGGAGGCTGAGGCAAGATGACTTGAACCCAGAGGCGCA 10848
QY 303 GAGGTTGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
Db 10847 GAGGTTGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10788
QY 363 AAATATGATTAATAAGAGCAAGGTGACCAAAAGAGATAGGCTGAAAAATTGT 422
Db 10787 CTGTCTCAAAAAAATAATTAATAAGTACACAAAAAATTAAAAATAATAAACACATGT 10728
QY 423 CTAAA 427
Db 10727 AGACA 10723
```

Search completed: April 18, 2006, 20:28:51
Job time : 257 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 18, 2006, 21:11:57 ; Search time 1115 Seconds

(without alignments)
8877.530 Million cell updates/sec

Title: US-10-089-928-1
Perfect score: 1197
Sequence: 1 ttctattcacacaaga.....cagggtcttctcatcagtt 1197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 segs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq: *
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	467	39.0	752	6	US-10-012-697-457
2	467	39.0	752	9	US-10-779-543-22457
3	441.2	36.9	554	5	US-10-027-632-86702
4	441.2	36.9	554	6	US-10-027-632-86702
5	186.6	15.6	36000	9	US-10-948-947A-27
6	186.2	15.6	865	4	US-09-925-065A-58507
7	184.2	15.4	324	8	US-10-674-124A-24268
8	182	15.2	47804	6	US-10-017-161-717
9	182	15.2	47804	6	US-10-292-798-629
10	179.8	15.0	1538	4	US-09-925-065A-685565
11	179.8	15.0	1538	4	US-09-925-065A-685566
12	176.6	14.8	412	8	US-10-674-124A-304
13	175	14.6	95832	3	US-10-741-600-17662
14	174.8	14.6	397658	3	US-09-813-320-3
15	174.8	14.6	397658	10	US-11-044-879-3
16	174.6	14.6	60463	8	US-10-719-993-6783
17	174	14.5	555	4	US-09-925-065A-164271
18	172.8	14.4	560	4	US-09-925-065A-164273
19	172.4	14.4	555	4	US-09-925-065A-164272
20	172	14.4	433	8	US-10-674-124A-13089
21	171.2	14.3	1601	4	US-09-925-065A-61806
22	171.2	14.3	378361	3	US-09-901-136-3
23	171.2	14.3	378361	8	US-10-483-329-3

C	24	171	14.3	610	4	US-09-925-065A-819115	Sequence 819115,
C	25	171	14.3	617	4	US-09-925-065A-837471	Sequence 837471,
C	26	171	14.3	1561	5	US-10-027-632-253907	Sequence 253907,
C	27	171	14.3	1561	6	US-10-027-632-253907	Sequence 253907,
C	28	170.6	14.3	1277	4	US-09-925-065A-74913	Sequence 74913, A
C	29	170.6	14.3	1561	5	US-10-027-632-253906	Sequence 253906,
C	30	170.6	14.3	1561	6	US-10-027-632-253906	Sequence 253906,
C	31	170.6	14.3	744802	6	US-10-292-798-1369	Sequence 1369, Ap
C	32	169.4	14.2	449	8	US-10-674-124A-25360	Sequence 25360, A
C	33	169.4	14.2	482	4	US-09-925-065A-436024	Sequence 436024,
C	34	169.4	14.2	482	4	US-09-925-065A-436025	Sequence 436025,
C	35	169.4	14.2	703	4	US-09-925-065A-942392	Sequence 942392,
C	36	169	14.1	482	4	US-09-925-065A-436023	Sequence 436023,
C	37	169	14.1	509	9	US-10-779-543-17664	Sequence 17664, A
C	38	168.2	14.1	574	4	US-09-925-065A-474895	Sequence 474895,
C	39	167.8	14.0	133100	9	US-10-723-681-4	Sequence 4, Appli
C	40	167	14.0	149382	8	US-10-741-600-17661	Sequence 17661, A
C	41	166.8	13.9	549	9	US-10-779-543-18645	Sequence 18645, A
C	42	166.4	13.9	530	4	US-09-925-065A-56157	Sequence 56157, A
C	43	166.4	13.9	23544	3	US-09-764-891-8517	Sequence 8517, Ap
C	44	166.2	13.9	454	8	US-10-674-124A-23373	Sequence 23373, A
C	45	165.8	13.9	6638	6	US-10-220-891-40	Sequence 40, Appli

ALIGNMENTS

RESULT 1
US-10-012-697-457
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Scott, Beth
; APPLICANT: Drmanac, Radoje
; APPLICANT: Cirkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
; FILE REFERENCE: 2300-16252
; CURRENT APPLICATION NUMBER: US/10/012,697
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 60/254,648
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 60/275,668
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 1568
; SOFTWARE: FastSeq for Windows Version 4.0
; OTHER INFORMATION: n = A,T,C or G
; US-10-012-697-457

Query Match 39.0%; Score 467; DB 6; Length 752;
Best Local Similarity 90.9%; Pred. No. 4,7e-115;
Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;

Qy	1	TTTCTATTTCACACAGCAAGAAAGATGAGAGAGAGCTAGAAAGTAGATGTGATC	60
Db	74	TTTCTATTTCACACAGCAAGCAAGAAAGATGAGAGAGAGCTAGAAAGTAGATGTGATC	133
Qy	61	ATATGAATATGATTTTCTGCTTTTGCATGTATGTGACACATGACAGAGTGAC	120
Db	134	ATATGAATATGATTTTCTGCTTTTGCATGTATGTGACACATGACAGAGTGAC	193
Qy	121	AGCAGAGTTGAGACGAGCTGACCAACATGATGAAATCCCGTCTACTAAACACACA	180

```
Db 194 AGCAGAGTTGAGACCAACCTGACCACATGTTGAATCCCGTCTTACTTAACACACA 253
QY 181 CACACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGG 240
Db 254 CACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGG 313
QY 241 GCACCTGTATCCAGCTACTTGGAGGCTGAGGACAAAGATGACTTGAACCCAGAGG 300
Db 314 GCACCTGTATCCAGCTACTTGGAGGCTGAGGACAAAGATGACTTGAACCCAGAGG 373
QY 301 CGGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGTGA 360
Db 374 CGGNGNTGCAGNAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGTGA 433
QY 361 AAAAATATGATATAAAGAGAGCAAGGTGACCAAAAGAGATAGGCTGAAAAATTT 420
Db 434 AAAAATAGTATATANAGAGAGCNGGGTGACCAAAAGAGATANGNTGAAANATTT 493
QY 421 GTCTAAATGTTG-CCTCTTCTTATAGCTGATATGTTAAGTTTATTTTCCCTAG 479
Db 494 GTCTAANNGGGGCCCTCTTCTTATAGCTGATATGNTAAGTTTATTTTCCCTA- 552
QY 480 TAGCGAATTCTAAGGATGAAGAGAATCCTTTTCAGTTTACTTCCCAAGTGTGTA 539
Db 553 NAGCGANTTCT-AGGGATGAANAAAAATNCTTTTCAAGTTTCTCCCAAGGNGTATA 611
QY 540 TAACTACTATAGTGAATATAAGTCCCAATTTATCTTTGAA 581
Db 612 CCTCTA---TGGGAATATAATAGCCCAATTTTTTGGAGTA 649
```

RESULT 2

```
US-10-779-543-22457
; Sequence 22457, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; PRIOR FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22457
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 33, 311, 322, 345, 363, 377, 380, 386, 408, 450, 458, 480,
; LOCATION: 482, 489, 500, 501, 533, 553, 559, 573, 581, 606, 666, 681,
```

```
; LOCATION: 686, 693, 697, 706, 709, 711, 721, 733, 736, 747, 751
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-22457
```

```
Query Match 39.0%; Score 467; DB 9; Length 752;
Best Local Similarity 90.9%; Pred. No. 4.7e-115;
Matches 529; Conservative 0; Mismatches 46; Indels 7; Gaps 4;
```

```
QY 1 TTTCTATTTCACAACAAGCAAGAAAAAAGATGAGAGAGGACTAGAAAGTAGATGTGATC 60
Db 74 TTTCTATTTCACAACAAGCAAGAAAAAAGATGAGAGAGGACTAGAAAGTAGATGTGATC 133
QY 61 ATATGAATATGATTTTCTTGTCTTTTGTGATGATGTGTGACATGACAGAGTGAC 120
Db 134 ATATGAATATGATTTTCTTGTCTTTTGTGATGATGTGTGACATGACAGAGTGAC 193
QY 121 AGCAGAGTTGAGACCAAGCCTGACCAATGTGTAATCCGCTCTACTAAACACACA 180
Db 194 AGCAGAGTTGAGACCAAGCCTGACCAATGTGTAATCCGCTCTACTAAACACACA 253
QY 181 CACACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGG 240
Db 254 CACACACACACACACACACACACACACACACACATAGCCGGGATGTGTGG 313
QY 241 GCACCTGTATCCCAAGCTACTTGGAGGCTGAGGACAAAGATGACTTGAACCCAGAGG 300
Db 314 GCACCTGTATCCCAAGCTACTTGGAGGCTGAGGACAAAGATGACTTGAACCCAGAGG 373
QY 301 CGGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGTGA 360
Db 374 CGGNGNTGCAGNAGCTGAGATCATGCCATTGCACTCCAGCCGTGGTGAAGTGA 433
QY 361 AAAAATATGATATAAAGAGAGCAAGGTGACCAAAAGAGATAGGCTGAAAAATTT 420
Db 434 AAAAATAGTATATANAGAGAGCNGGGTGACCAAAAGAGATANGNTGAAANATTT 493
QY 421 GTCTAAATGTTG-CCTCTTCTTATAGCTGATATGTTAAGTTTATTTTCCCTAG 479
Db 494 GTCTAANNGGGGCCCTCTTCTTATAGCTGATATGNTAAGTTTATTTTCCCTA- 552
QY 480 TAGCGAATTCTAAGGATGAAGAGAATCCTTTTCAGTTTACTTCCCAAGTGTGTA 539
Db 553 NAGCGANTTCT-AGGGATGAANAAAAATNCTTTTCAAGTTTCTCCCAAGGNGTATA 611
QY 540 TAACTACTATAGTGAATATAAGTCCCAATTTATCTTTGAA 581
Db 612 CCTCTA---TGGGAATATAATAGCCCAATTTTTTGGAGTA 649
```

RESULT 3

```
US-10-027-632-86702/c
; Sequence 86702, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```


;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 86702
;; LENGTH: 554
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-86702

Query Match 36.9%; Score 441.2; DB 5; Length 554;
Best Local Similarity 96.5%; Pred. No. 3.6e-108;
Matches 473; Conservative 1; Mismatches 9; Indels 7; Gaps 2;

```
QY 1 TTTCTATTTCACCAAGCAAGCAAAAAAGATGAGAGAGAGCTAGAAAGTAGATGTGATC 60
    |||||||
Db 484 TTCTATTCTCACCAGCAGCAAGAAAAAGATGAGAGAGAGCTAGAAAGTAGATGTGATC 425
    |||||||
QY 61 ATATGAATATGATTTCTCTGCTTTTTCATGTATGTGTGACACATGCAGAAAGTGAC 120
    |||||||
Db 424 ATATGAATATGATTTCTCTGCTTTTTCATGTATGTGTGACACATGCAGAAAGTGAC 365
    |||||||
QY 121 AGCAGAGTTCGAGACCAAGCCTGACCAACATGTGTAAATCCCGTCTACTTAACACACA 180
    |||||||
Db 364 AGCAGAGTTCGAGACCAAGCCTGACCAACATGTGTAAATCCCGTCTACTTAACACA 311
    |||||||
QY 181 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 240
    |||||||
Db 310 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 251
    |||||||
QY 241 GCACCTGTAATCCCAAGCTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAGAGG 300
    |||||||
Db 250 GCACCTGTAATCCCAAGCTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAGAGG 191
    |||||||
QY 301 CGAGGTTGACGTGAGCTGAGATCATGTCATGCTCCAGCCTGGTGACGATGAA 360
    |||||||
Db 190 CGAGGTTGACGTGAGCTGAGATCATGTCATGCTCCAGCCTGGTGACGATGAA 131
    |||||||
QY 361 AAAAATAATGATAATAAAGAGCAAGGTGACCAAAAGAGATAGCTGAAAAATTT 420
    |||||||
Db 130 AAAAATAATGATAATAAAGAGCAAGGTGACCAAAAGAGATAGCTGAAAAATTT 71
    |||||||
QY 421 GTCTAATGTGCGCTCTTCTCTTATAGCTGATATGTTAAGTTATTTT-CCCTAG 479
    |||||||
Db 70 GTCTAATGTGCGCTCTTCTCTTATAGCTGATATGTTAAGTTATTTTCCCTAG 11
    |||||||
QY 480 TAGCGAATTC 489
    |||||||
Db 10 TAGCGAATTC 1
```

RESULT 4

US-10-027-632-86702/c
; Sequence 86702, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 86702
;; LENGTH: 554
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-86702

Query Match 36.9%; Score 441.2; DB 6; Length 554;
Best Local Similarity 96.5%; Pred. No. 3.6e-108;
Matches 473; Conservative 1; Mismatches 9; Indels 7; Gaps 2;

```
QY 1 TTTCTATTTCACCAAGCAAGCAAAAAAGATGAGAGAGAGCTAGAAAGTAGATGTGATC 60
    |||||||
Db 484 TTCTATTCTCACCAGCAGCAAGAAAAAGATGAGAGAGAGCTAGAAAGTAGATGTGATC 425
    |||||||
QY 61 ATATGAATATGATTTCTCTGCTTTTTCATGTATGTGTGACACATGCAGAAAGTGAC 120
    |||||||
Db 424 ATATGAATATGATTTCTCTGCTTTTTCATGTATGTGTGACACATGCAGAAAGTGAC 365
    |||||||
QY 121 AGCAGAGTTCGAGACCAAGCCTGACCAACATGTGTAAATCCCGTCTACTTAACACACA 180
    |||||||
Db 364 AGCAGAGTTCGAGACCAAGCCTGACCAACATGTGTAAATCCCGTCTACTTAACACA 311
    |||||||
QY 181 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 240
    |||||||
Db 310 CACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGTGTGG 251
    |||||||
QY 241 GCACCTGTAATCCCAAGCTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAGAGG 300
    |||||||
Db 250 GCACCTGTAATCCCAAGCTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAGAGG 191
    |||||||
QY 301 CGAGGTTGACGTGAGCTGAGATCATGTCATGCTCCAGCCTGGTGACGATGAA 360
    |||||||
Db 190 CGAGGTTGACGTGAGCTGAGATCATGTCATGCTCCAGCCTGGTGACGATGAA 131
    |||||||
QY 361 AAAAATAATGATAATAAAGAGCAAGGTGACCAAAAGAGATAGCTGAAAAATTT 420
    |||||||
Db 130 AAAAATAATGATAATAAAGAGCAAGGTGACCAAAAGAGATAGCTGAAAAATTT 71
    |||||||
QY 421 GTCTAATGTGCGCTCTTCTCTTATAGCTGATATGTTAAGTTATTTT-CCCTAG 479
    |||||||
Db 70 GTCTAATGTGCGCTCTTCTCTTATAGCTGATATGTTAAGTTATTTTCCCTAG 11
    |||||||
QY 480 TAGCGAATTC 489
    |||||||
Db 10 TAGCGAATTC 1
```

RESULT 5

US-10-948-947A-27/c
; Sequence 27, Application US/10948947A
; Publication No. US20050130924A1
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Freier, Susan M.
; APPLICANT: Manoharan, Muthiah
; APPLICANT: Gaarde, William A.
; APPLICANT: Griffey, Richard H.
; APPLICANT: Swayze, Eric B.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: ANTISENSE INHIBITION VIA RNASE H-INDEPENDENT REDUCTION IN mRNA
; FILE REFERENCE: ISPH-0871
; CURRENT APPLICATION NUMBER: US/10/948,947A
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: 60/392,020
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 10/461,163
; PRIOR FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: FastSeq for Windows Version 4.0

Db 145 AGCGGAGCTTGACGTGACCCGAGATCATGCCACTGCACTCCAGCCTGGTGACACAGCA 86
QY 358 AAAAAAAAAATGATTAATAAAGAGAGCAAGTGACCAAAA 398
Db 85 AGACTCGGTCTCAAAAAAAAACAAAACAAAACAAAACAAA 45

RESULT 8

US-10-017-161-717
; Sequence 717, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT FILING DATE: 2002-12-18
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 717
; LENGTH: 47804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(47804)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1022)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1093)..(1208)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1871)..(2040)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3227)..(3386)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3492)..(3698)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5159)..(5424)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5651)..(6504)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18486)..(18612)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19045)..(19140)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25901)..(26126)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26658)..(26848)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30445)..(30482)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37501)..(37634)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (38681)..(38889)

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46474)..(46682)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46867)..(47604)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (7478)..(7577)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (17033)..(17132)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (24561)..(24660)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (33542)..(33641)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (41424)..(41523)
; OTHER INFORMATION: a, t, c, g, unknown or other
; US-10-017-161-717

Query Match 15.2%; Score 182; DB 6; Length 47804;
Best Local Similarity 82.2%; Pred. No. 2.4e-37;
Matches 222; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 118 GACAGCAGAGTTGAGAGCCAGCTGACCAACATGTTGAATCCGCTCTACTAA--A 174
Db 12662 GAGTTCAGAGTTGAGACTAGCTGGCCAACATAGCAAAACCCATCTCTACTAAACA 12721
QY 175 CACACACACACACACACACACACACACACACACACACACACACAAATAGCCGGCATGG 234
Db 12722 CACACACACACACACACACACACACACACACACACACACACACATGGCCAGTATGA 12781
QY 235 TGGTGGCAGCTGTAATCCAGCTACTGGAGGCTGAGGCAAGAATGACTGAACCC 294
Db 12782 TGGTGTGACCTGTAATCCAGCTACTGAGGAGCTGAGGCAAGATCACTGAACCC 12841
QY 295 AGAGGCGGAGGTTGAGTGAGCTGAGATCATGCCATTGCACCTCGGTGACGAG 354
Db 12842 GGGAGGTGAGGTTGAGTGAGCTGAGATTTGCCATCAGCTCCAGCCTGGCGACAAG 12901
QY 355 TGAATAAAAAAATATGATTAATAAAGAGAC 384
Db 12902 AGCGAACTCCGTCTCAAAAAAAAAAAC 12931

RESULT 9

US-10-292-798-629
; Sequence 629, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT FILING DATE: 2002-11-13
; CURRENT APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 629


```
LENGTH: 47804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(47804)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(1022)
FEATURE:
NAME/KEY: CDS
LOCATION: (1093)..(1208)
FEATURE:
NAME/KEY: CDS
LOCATION: (1871)..(2040)
FEATURE:
NAME/KEY: CDS
LOCATION: (3227)..(3386)
FEATURE:
NAME/KEY: CDS
LOCATION: (3492)..(3698)
FEATURE:
NAME/KEY: CDS
LOCATION: (5159)..(5424)
FEATURE:
NAME/KEY: CDS
LOCATION: (5651)..(6504)
FEATURE:
NAME/KEY: CDS
LOCATION: (18486)..(18612)
FEATURE:
NAME/KEY: CDS
LOCATION: (19045)..(19140)
FEATURE:
NAME/KEY: CDS
LOCATION: (25901)..(26126)
FEATURE:
NAME/KEY: CDS
LOCATION: (26658)..(26848)
FEATURE:
NAME/KEY: CDS
LOCATION: (30445)..(30482)
FEATURE:
NAME/KEY: CDS
LOCATION: (37501)..(37634)
FEATURE:
NAME/KEY: CDS
LOCATION: (38681)..(38889)
FEATURE:
NAME/KEY: CDS
LOCATION: (46474)..(46682)
FEATURE:
NAME/KEY: CDS
LOCATION: (46867)..(47604)
FEATURE:
NAME/KEY: modified_base
LOCATION: (7478)..(7577)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (17033)..(17132)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (24561)..(24660)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (33542)..(33641)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
```

```
LOCATION: (41424)..(41523)
OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-629
```

```
Query Match 15.2%; Score 182; DB 6; Length 47804;
Best Local Similarity 82.2%; Pred. No. 2.4e-37;
Matches 222; Conservative 0; Mismatches 45; Indels 3; Gaps 1;
```

```
QY 118 GACAGCAGGAGTTGAGACCAAGCTGACCAACATGTTGAAATCCGCTCTACTAA--A 174
DB 12662 GAGTCAGGAGTTGAGACTAGCTGGCCAAACATAGCAAAACCCATCTCTACTAAACA 12721
```

```
QY 175 CACACACACACACACACACACACACACACACACACACACACACACATAGCCGGCATGG 234
DB 12722 CACACACACACACACACACACACACACACACACACACACACACATTGGCCAGTATGA 12781
```

```
QY 235 TGGTGGCACCTGTAATCCAGCTACTTGGAGGCTGAGCCACAAGATGACTTGAACCC 294
DB 12782 TGGTGTGACACCTGTAATCCAGCTACTGAGAGGCTGAGGACAGAGATCACTTGAACCC 12841
```

```
QY 295 AGAGCGGAGGTTGACAGTGAATCATGATCATGCACTCCAGCCTGGGTGACGAG 354
DB 12842 GGGAGGTGAGGTTGACAGTGAATCATGATGTTGTCATCACTCCAGCCTGGGCAAG 12901
```

```
QY 355 TGAATAAATAATATGATTAATAAGAGAGC 384
DB 12902 AGCGAACTCCGTCTCAAAAAAACC 12931
```

RESULT 10

```
US-09-925-065A-685565/c
Sequence 685565, Application US/09925065A
Publication No. US20050228172A9
```

GENERAL INFORMATION:

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 685565
LENGTH: 1538
```

```
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-685565
```

```
Query Match 15.0%; Score 179.8; DB 4; Length 1538;
Best Local Similarity 85.5%; Pred. No. 1.7e-37;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1;
```

```
QY 117 TGACAGCAGGAGTTGAGACCAAGCTGACCAACATGTTGAAATCCGCTCTACTAA--- 173
DB 423 TGAGGTGAGGAGTTGAGACCAAGCTGACCAACATGTTGAAATCCGCTCTACTAAAT 364
```

```
QY 174 -ACACACACACACACACACACACACACACACACACACACACACACAATAGCCGGCAT 232
DB 363 TACACACACACACACACACACACACACACACACACACACACACAATAGCTGGCGT 304
```

```
QY 233 GGTGTGGGACCTGTAATCCAGCTACTTGGGAGGCTGAGGCAAGATGACTTGAAC 292
DB 303 GGTGTGGGACCTGTAATCCAGCTACTTAGGAGGCTGAGGCAAGATGCTTGAAC 244
```

QY 293 CCAGAGCGGAGTTGACGTGACATGATCATGCTTGCCTCCAGCCTGGGTGACG 352
Db 243 CTGGAGGTGAGGTTGACGTGAGCCAGATCATGCTGCACTCCAGCCTAGGCAACA 184
QY 353 AGTGAAAA 361
Db 183 AGAGGGAGA 175

RESULT 11

US-09-925-065A-685566/c
; Sequence 685566, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 685566
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-685566

Query Match 15.0%; Score 179.8; DB 4; Length 1538;
Best Local Similarity 85.5%; Pred. No. 1.7e-37;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1;

QY 117 TGACAGCAGGAGTTGAGACCAAGCTGACCAATGCTGAAATCCGCTCTACTAA--- 173
Db 423 TGAGGTAGGAGTTGAGACCAAGCTGACCAATGCTGAAATCCGCTCTACTAA--- 364
QY 174 -ACACACACACACACACACACACACACACACACACACACACAAATAGCGGCAT 232
Db 363 TACACACACACACACACACACACACACACACACACACACAAATAGCTGGCGT 304
QY 233 GGTGGGGCACCTGTAATCCAGTACTTGGAGGCTGAGGCAAGAATGACTTGAAC 292
Db 303 GGTGGGGCACCTGTAATCCAGTACTTGGAGGCTGAGGCAAGAATGACTTGAAC 244
QY 293 CCAGAGCGGAGGTTGACGTGACATGATCATGCTTGCCTCCAGCCTGGGTGACG 352
Db 243 CTGGAGGTGAGGTTGACGTGAGCCAGATCATGCTCCAGCCTAGGCAACA 184
QY 353 AGTGAAAA 361
Db 183 AGAGGGAGA 175

RESULT 12

US-10-674-124A-304
; Sequence 304, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMAYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS

; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 304
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AL031602.14_2459
; FEATURE:
; OTHER INFORMATION: Located on chromosome 1
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 48737
US-10-674-124A-304

Query Match 14.8%; Score 176.6; DB 8; Length 412;
Best Local Similarity 87.2%; Pred. No. 6.4e-37;
Matches 205; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 118 GACAGCAGGAGTTGAGACCAAGCTGACCAATGCTGAAATCCGCTCTACTAAACAC 177
Db 109 GAGGTAGGAGTTGAGACCAATGCTGACCAATGCTGAAATCCGCTCTACTAAACAC 168
QY 178 ACACACACACACACACACACACACACACACACACACACACAAATAGCGGCATG 236
Db 169 ACACACACACACACACACACACACACACACACACACACAAATAGCGGCATG 228
QY 237 GTGGCAGCTGTAATCCAGTACTTGGAGGCTGAGGCAAGAATGACTTGAACCCAG 296
Db 229 GCGGGCGCTGTAGTCCAGTACTTGGAGGCTGAGGCAAGAATGCGTGAACCCGG 288
QY 297 GAGGCGAGGTTGACGTGAGTGCATGATCATGCTTGCCTCCAGCCTGGGTGAC 351
Db 289 GAGGTGAGGTTGACGTGAGTGCATGATCATGCTTGCCTCCAGCCTGGGTGAC 343

RESULT 13

US-10-741-600-17662
; Sequence 17662, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17662
; LENGTH: 95832
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17662

Query Match 14.6%; Score 175; DB 8; Length 95832;

Best Local Similarity 78.0%; Pred. No. 2.6e-35;
Matches 224; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

[illegible]

RESULT 14

```

; Sequence 3, Application US/09813320
; Patent No. US20020142378A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, Hongyu et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001172
; CURRENT APPLICATION NUMBER: US/09/813,320
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 397658
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(397658)
; OTHER INFORMATION: n = A,T,C or G
US-09-813-320-3

```

Query Match	14.6%	Score 174.8;	DB 3;	Length 397658;
Best Local Similarity	80.4%;	Pred. No. 6e-35;		
Matches 217;	Conservative	0;	Mismatches 52;	Indels 1;
				Gaps 1;

OY	118	GACAGCAGGAGTTGCAGACCAGCTGACCAACATGGTGAATCCCGTCTCTACTTAACAC	177.
Dd	20400	GAGGTCAAGAGATGGAGACCATTCTGGCTTAACACAGTGAAACCCCGTCTCTACTTAAGAAC	20459
OY	178	ACACACACACACACACACACACACACACACACACACAC-ACAATAGCCGGCGCATGGTG	236
Dd	20460	ACACACACACACACACACACACACACACACACACACACAGATTAGCCGGCGCTGGTG	20519
OY	237	GTGGG ^a CCTGTAATCCCAGCTA ^c TGGAGGCTGAGGCCAAGAATGACTTGAACCCAG	296
Dd	20520	CTGGGGCCTGTAATCCCAGCTACTCCTGAGGCTGAGGCAGAGATGGCGTGAACCCAG	20579
OY	297	GAGGCGGAGGTTGCAGTGA ^c GTGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTG	356
Dd	20580	GAGGCGGAATTTCAGTGA ^c GTGATGATGAGGCCACTGCACTCCAGCCTGGGCGCACAGGC	20639
OY	357	AAAAAAAAATAATGATTAATAAGAGCAA	386
Dd	20640	GAGACTCCA ^c CTCAAAAAAAAAAAAAAAA	20669

RESULT 15
US-11-044-879-3

```

; Sequence 3, Application US/11044879
; Publication No. US20050130218A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, Hongyu et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001172CON
; CURRENT APPLICATION NUMBER: US/11/044,879
; CURRENT FILING DATE: 2005-01-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 397658
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(397658)
; OTHER INFORMATION: n = A,T,C or G
US-11-044-879-3

```

Query Match

	Matches	217;	Conservative	0;	Mismatches	52;	Indels	1;	Gaps	1;
QY	118	GACAGCAGGAGTTGAGAGCCAGCTGACCAACATGTTGAAATCCCGTCTCTACTTAACAC	177							
Db	20400	GAGGTACGAGATGAGACCATCCTGGCTTAACACAGTGAACCCCGTCTCTACTTAAGAC	20459							
QY	178	ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC	236							
Db	20460	ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC	20519							
QY	237	GTGGGCACCTGTAATCCCACTACTTGGGAGGCTGAGGCACAAGAATGACTTGAACCCAG	296							
Db	20520	CTGGGCGCTGTAATCCCACTACTTCTGAGGCTGAGGCAGGAAGATGGCGTGAACCCAG	20579							
QY	297	GAGCGCAGGTTGACGTGAGCTGAGATCATGTCATTGCACTCCAGCCTGGTGACGAGTG	356							
Db	20580	GAGCGGAATTGTCAGTGAAGTGAAGATAGCGCCACTGCACTCCAGCCTGGCGACAGGCGC	20639							
QY	357	AAAAAAAAAATATGATTAATATAAGAGCAA	386							
Db	20640	GAGACTCCACCTCAAAAAAAAAAAAAAAAAA	20669							

Search completed: April 18, 2006, 21:30:41
Job time : 1117 secs

Db 453 CACACACACACACACACACACACACACACACACACAAATAGCCGGGCATGTTGTGG 512
Qy 241 GCACCTGTAATCCAGCTACTTTGGAGGCTGAGGCACAAGAAATGACTTGAACCCAGAGG 300
Db 513 GCACCTGTAATCCAGCTACTTTGGAGGCTGAGGCACAAGAAATGACTTGAACCCAGAGG 572
Qy 301 CGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA AAA 360
Db 573 CGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA AAA 632
Qy 361 AAAAATAATGATAATAAAGAGAGCAAGGTGACCAAAAAGAGAAATAGGCTGAAAAATTT 420
Db 633 AAAAATAATGATAATAAAGAGAGCAAGGTGACCAAAAAGAGAAATAGGCTGAAAAATTT 692
Qy 421 GTCTAAATGGTGCCCTCTTCTTATAGCTGCATATGTTAAGTTATTTTCCCTAGT 480
Db 693 GTCTAAATGGTGCCCTCTTCTTATAGCTGCATATGTTAAGTTATTTTCCCTAGT 752
Qy 481 AGCGAATTCTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 540
Db 753 AGCGAATTCTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 812
Qy 541 AACTACTATAGTGAATAATAAGTCCAATTATCTTTGAAGTATAGTTAATATGTAACG 600
Db 813 AACTACTATAGTGAATAATAAGTCCAATTATCTTTGAAGTATAGTTAATATGTAACG 872
Qy 601 AACTCTTAAGGCCAGTTGTATACCCAGGCCAAACGCCCTTCTAACATCT 649
Db 873 AACTCTTAAGGCCAGTTGTATACCCAGGCCAAACGCCCTTCTAACATCT 921

RESULT 2
US-10-480-1157494
; Sequence 1157494, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1157494
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1157494

Query Match 54.2%; Score 649; DB 10; Length 921;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTCTATTTCACAACAAGCAAGAAAAAGATGAGAGGACTAGAAAAGTAGATGTATC 60
Db 273 TTCTATTTCACAACAAGCAAGAAAAAGATGAGAGGACTAGAAAAGTAGATGTATC 332
Qy 61 AATAGATAATGATTTCTGCTTTTTCATGTATGTGTGACACATGACAGAGTAC 120
Db 333 AATAGATAATGATTTCTGCTTTTTCATGTATGTGTGACACATGACAGAGTAC 392
Qy 121 AGCAGAGTTGAGAGCAGCCTGACCAACATGTGAATCCGCTCTACTAAACACACA 180
Db 393 AGCAGAGTTGAGAGCAGCCTGACCAACATGTGAATCCGCTCTACTAAACACACA 452
Qy 181 CACACACACACACACACACACACACACACACACATAAGCCGGGCATGTGTG 240
Db 453 CACACACACACACACACACACACACACACACACATAAGCCGGGCATGTGTG 512

Qy 241 GCACCTGTAATCCAGCTACTTTGGAGGCTGAGGCACAAGAAATGACTTGAACCCAGAGG 300
Db 513 GCACCTGTAATCCAGCTACTTTGGAGGCTGAGGCACAAGAAATGACTTGAACCCAGAGG 572
Qy 301 CGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA AAA 360
Db 573 CGAGGTTGCAGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA AAA 632
Qy 361 AAAAATAATGATAATAAAGAGAGCAAGGTGACCAAAAAGAGAAATAGGCTGAAAAATTT 420
Db 633 AAAAATAATGATAATAAAGAGAGCAAGGTGACCAAAAAGAGAAATAGGCTGAAAAATTT 692
Qy 421 GTCTAAATGGTGCCCTCTTCTTATAGCTGCATATGTTAAGTTATTTTCCCTAGT 480
Db 693 GTCTAAATGGTGCCCTCTTCTTATAGCTGCATATGTTAAGTTATTTTCCCTAGT 752
Qy 481 AGCGAATTCTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 540
Db 753 AGCGAATTCTAAGGATGAAGAAGAAATCCTTTTCAGTTTACTTCCCAAGGTGTAT 812
Qy 541 AACTACTATAGTGAATAATAAGTCCAATTATCTTTGAAGTATAGTTAATATGTAACG 600
Db 813 AACTACTATAGTGAATAATAAGTCCAATTATCTTTGAAGTATAGTTAATATGTAACG 872
Qy 601 AACTCTTAAGGCCAGTTGTATACCCAGGCCAAACGCCCTTCTAACATCT 649
Db 873 AACTCTTAAGGCCAGTTGTATACCCAGGCCAAACGCCCTTCTAACATCT 921

RESULT 3
US-10-301-480-576225/c
; Sequence 576225, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 576225
; LENGTH: 989
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-576225

Query Match 15.6%; Score 186.6; DB 10; Length 989;
Best Local Similarity 77.9%; Pred. No. 2.1e+02;
Matches 225; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy 118 GACAGCAGAGTTGAGACCAAGCCTGACCAACATGGTGAATCCCGTCTACTAAACAC 177
Db 554 GAGATCAGAGATCGAGACCATCCCGCTAACATGTGAACCTGTCTACTAAAAAT 495
Qy 178 ACAACACACACACACACACACACACACACACACACAATAGCCGGGCATGTG 237
Db 494 ACAACACACACACACACACACACACACACACACACAATAGCCGGGCATGTG 435
Qy 238 TGGGCACTGTATCCCACTACTTTGGAGGCTGAGGCACAAGATGACTTGAACCAAG 297
Db 434 CAGGCACTGTAGCCCACTACTTTGGAGGCTGAGGCACAAGATGACTTGAACCAAG 375
Qy 298 AGCGGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 357
Db 374 AGCGGAGGTTGACGTGAGCTGAGATCATGCCATTGCACTCCAGCCTGGGTGACGAGTGA 315

QY 298 AGCGGAGGTTGCAGTGAGCTGAGATCATGCGCATTTGCACCTCCAGCCTGGGTGACGAGTGA 357
 |||||
 Db 269 AGCGGAGCTTGCAGTGAGCCAGATCACGCCACTGCACTTAAGCCTGGGCGGACGAGTGT 210
 |||||
 QY 358 AAAAAAAAAATGATTAATAAAGAGAGCAAGTGCACCAAAAAGAGATA 406
 |||||
 Db 209 AGACTCTGTCTCAAAAAAAAAAAAAAAAACAAAAAAAAACAAAA 161
 |||||

```

RESULT 7
US-10-301-480-773154/c
; Sequence 773154, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 773154
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-773154

```

[illegible]

```

RESULT 8
US-10-995-561-13274/c
; Sequence 13274, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13274
;
; LENGTH: 415117

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (415117)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables :
US-10-995-561-13274

```

Query Match	15.4%;	Score 184;	DB 8;	Length 415117;
Best Local Similarity	87.3%;	Pred. No. 3.6;		
Matches 213;	Conservative 0;	Mismatches 30;	Indels 1;	Gaps 1;
QY 118	GACAGCAGAGATTGAGAGACCAGCCTGACCAACATGCTGAATCCCGTCTTACTTAACAC	177		
Db 342008	GAGCTCAGAGATTGAGAGCCAGCCTGGGCAACA-GCCGAAACCTGCTCTTACTTAATAAT	341950		
QY 178	ACACACACACACACACACACACACACACACACACACACACACACACTAGCCAGGATGCTGG	237		
Db 341949	ACACACACACACACACACACACACACACACACACACACACACACTAGCCAGGATGCTGG	341890		
QY 238	TGGGCACCTGTAAATCCAGCTACTTGGAGGCTGAGGCACAAGAATGACTTGAACCCAGG	297		
Db 341889	CGCACGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCACAAGAATGCTTGAACCCGGG	341830		
QY 298	AGCGGAGGTTGCACTGAGCTGAGATCATGCCATTGCCACTCCAGCCTGGGTGACGAGTGA	357		
Db 341829	AGGTGAGGTTTCACTGAGCCGAGATCATGCCACTCCAGCCTGGGCAACAAGAGA	341770		
QY 358	AAAA 361			
Db 341769	GAAA 341766			

```

RESULT 9
US-10-301-480-577420/c
; Sequence 577420, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577420
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-577420

```

Query Match	15.1%;	Score 180.2;	DB 10;	Length 754;
Best Local Similarity	85.8%;	Pred. No. 3.8e+02;		
Matches 212;	Conservative	0;	Mismatches 33;	Indels 2;
			Gaps	1;
Oy	117	TCACAGCAGGAGTTCGAGACCAGCCTGACCAACATGSTGAATCCCGTCTCTACTTAAC-	175	
Db	434	TGAGGTCAAGAGTTTGAGACCAGCCTGGCCAAACATGSTGAANAACCCCTTCTTAACATAAAT	375	
Oy	176	-ACACACACACACACACACACACACACACACACACACACACACACAAATATAGCCGGCATGG	234	
Db	374	TACACACACACACACACACACACACACACACACACACACACACAAATTAGCTGGGCGGTGG	315	
Oy	235	TGGTGGGCACCTGTAATCCCAAGCTACTTTGGGAGGCTGAGGCAACAAGATGACTTGAACCC	294	
Db	314	TGGTGGGCACCTGTAATCCCAAGCTACTTAGGAGGCTGAGGCAAGAAATTGCTTGAACCT	255	
Oy	295	AGGAGCGGAGGTTGCAGTGAAGCTGAGATCATGCAATTGCACCTCCAGCCTGGGTGACGAG	354	

```
Db      254 GGGAGGTGAGGTTGACGTGAGCCAGATCATGCCATGCGCACTCCAGCCTTAGGCAACAAG 195
QY      355 TGA AAA 361
Db      194 AGGAGAGA 188
```

RESULT 10

```
US-10-301-480-1190829/c
; Sequence 1190829, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1190829
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1190829
```

```
Query Match      15.1%; Score 180.2; DB 10; Length 754;
Best Local Similarity 85.8%; Pred. No. 3.8e+02;
Matches 212; Conservative 0; Mismatches 33; Indels 2; Gaps 1;
```

```
QY      117 TGACAGCAGAGTTGAGACCGCTGACCAACATGTGAATCCGCTCTACTAAC- 175
Db      434 TGAGGTGAGAGTTTGAGACCGCTGAGCAACATGTGAATCCCTTCTTAATAAT 375
QY      176 -ACACACACACACACACACACACACACACACACACACACATAAGCCGGCATGG 234
Db      374 TACACACACACACACACACACACACACACACACACACAAATTAGCTGGGGCTGG 315
QY      235 TGGTGGGCACTGTAATCCAGCTACTTGGGAGGCTGAGGCACAGAATGACTTGAACCC 294
Db      314 TGGTGGGCACTGTAATCCAGCTACTTGAAGGCTGAGGCAGAGAATTGCTTGAACCT 255
QY      295 AGGAGCGGAGGTGACGTGAGCTGATCATGCCATTGCACTCCAGCCTGGGTGACGAG 354
Db      254 GGGAGGTGAGGTTGACGTGAGCCAGATCATGCCATGCACTCCAGCCTAGGCAACAAG 195
QY      355 TGA AAA 361
Db      194 AGGAGAGA 188
```

RESULT 11

```
US-09-925-065A-685565/c
; Sequence 685565, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
```

```
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 685565
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-685565
```

```
Query Match      15.0%; Score 179.8; DB 6; Length 1538;
Best Local Similarity 85.5%; Pred. No. 2.3e+02;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1;
```

```
QY      117 TGACAGCAGAGTTGAGACCGCTGACCAACATGTGAATCCGCTCTACTAA--- 173
Db      423 TGAGGTGAGAGTTTGAGACCGCTGAGCAACATGTGAATCCCTTCTTAATAAAT 364
QY      174 -ACACACACACACACACACACACACACACACACACACACATAAGCCGGCAT 232
Db      363 TACACACACACACACACACACACACACACACACACAAATTAAGCTGGCGT 304
QY      233 GGTGTGGGCACTGTAATCCAGCTACTTGGGAGGCTGAGGCACAGAATGACTTGAAC 292
Db      303 GGTGTGGGCACTGTAATCCAGCTACTTGAAGGCTGAGGCAGAGAATGCTTGAAC 244
QY      293 CCAGAGCGGAGGTTGACGTGAGCTGATCATGCCATTGCACTCCAGCCTGGTGACG 352
Db      243 CTGGAGGTGAGGTTGACGTGAGCCAGATCATGCCATGGCACTCCAGCCTAGGCAACA 184
QY      353 AGTGA AAA 361
Db      183 AGGAGAGA 175
```

RESULT 12

```
US-09-925-065A-685566/c
; Sequence 685566, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 685566
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-685566
```

```
Query Match      15.0%; Score 179.8; DB 6; Length 1538;
Best Local Similarity 85.5%; Pred. No. 2.3e+02;
Matches 213; Conservative 0; Mismatches 32; Indels 4; Gaps 1;
```

```
QY      117 TGACAGCAGAGTTGAGACCGCTGACCAACATGTGAATCCGCTCTACTAA--- 173
Db      423 TGAGGTGAGAGTTTGAGACCGCTGAGCAACATGTGAATCCCTTCTTAATAAAT 364
```

[illegible]

```

RESULT 13
US-10-301-480-542985/C
; Sequence 542985, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 542985
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-542985

```

[illegible]

```

RESULT 14
US-10-301-480-1156394/c
; Sequence 1156394, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

```

```

; TITLE OF INVENTION:  in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1156394
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1156394

```

[illegible]

```

RESULT 15
US-10-301-480-604426
; Sequence 604426, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 604426
;
; LENGTH: 904
;
; TYPE: DNA
;
; ORGANISM: Homo sapien
US-10-301-480-604426

```

Query Match	14.6%;	Score 174.6;	DB 10;	Length 904;
Best Local Similarity	86.2%;	Pred. No. 4.6e+02;		
Matches 206;	Conservative	0;	Mismatches 29;	Indels 4;
			Gaps	1;
Qy	117	TGACAGCAGAGATTGAGAGACCAGCCTGACCAACATGTGTGAATCCCCGCTCTACTAA---	173	
Db	392	TGAGGTGAGGAGTTCAAGACCAGCCTGACCAACATGTGTGAACCACAGTCTCTACTAACA	451	
Qy	174	-ACACACACACACACACACACACACACACACACACACACACCAATAGCCGGGCAT	232	

